

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 843 782**

51 Int. Cl.:

H04N 21/854 (2011.01)

H04N 21/236 (2011.01)

H04N 21/434 (2011.01)

H04N 21/435 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.01.2018 PCT/US2018/012476**

87 Fecha y número de publicación internacional: **12.07.2018 WO18129245**

96 Fecha de presentación y número de la solicitud europea: **05.01.2018 E 18709128 (5)**

97 Fecha y número de publicación de la concesión europea: **14.10.2020 EP 3566458**

54 Título: **Diseño de esquema restringido mejorado para vídeo**

30 Prioridad:

09.01.2017 US 201762444285 P
04.01.2018 US 201815862046

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
20.07.2021

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)
5775 Morehouse Drive
San Diego, CA 92121-1714, US

72 Inventor/es:

WANG, YEKUI

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 843 782 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Diseño de esquema restringido mejorado para vídeo

5 CAMPO

[0001] Esta solicitud está relacionada con esquemas restringidos para vídeo. Más específicamente, esta solicitud se refiere a sistemas y procedimientos para proporcionar un diseño mejorado del esquema restringido para un formato de archivo de medios.

10 ANTECEDENTES

[0002] Muchos dispositivos y sistemas permiten que los datos de vídeo sean procesados y enviados para su consumo. Los datos de vídeo digital incluyen grandes cantidades de datos para satisfacer las demandas de los consumidores y proveedores de vídeo. Por ejemplo, los consumidores de datos de vídeo desean vídeo de la máxima calidad, con alta fidelidad, resoluciones, velocidades de trama y similares. Como resultado, la gran cantidad de datos de vídeo que se requieren para satisfacer estas demandas supone una carga para las redes y dispositivos de comunicación que procesan y almacenan los datos de vídeo.

[0003] Se pueden usar diversas técnicas de codificación de vídeo para comprimir datos de vídeo. La codificación de vídeo se realiza de acuerdo con una o más normas de codificación de vídeo. Por ejemplo, los estándares de codificación de vídeo incluyen codificación de vídeo de alta eficiencia (HEVC), codificación de vídeo avanzada (AVC), codificación de grupo de expertos en imágenes en movimiento (MPEG), o similares. La codificación de vídeo utiliza, en general, procedimientos de predicción (por ejemplo, interpredicción, intrapredicción o similares) que aprovechan la redundancia presente en las imágenes o secuencias de vídeo. Un objetivo importante de las técnicas de codificación de vídeo es comprimir los datos de vídeo en una forma que use una velocidad de transferencia de bits más baja, mientras se evitan o minimizan las degradaciones en la calidad de vídeo. Con la disponibilidad de servicios de vídeo en constante evolución, se necesitan técnicas de codificación con una mejor eficacia de codificación.

30 BREVE EXPLICACIÓN

[0004] En algunos ejemplos, se describen en el presente documento técnicas y sistemas para proporcionar un diseño mejorado del esquema restringido para un formato de archivo de medios. El formato de archivo de medios puede incluir el formato de archivo de medios de base ISO (ISOBMFF) u otro formato de archivo de medios adecuado. ISOBMFF es un ejemplo de formato de archivo de medios, que incluye un diseño de esquema restringido. El diseño de esquema restringido en ISOBMFF se utiliza para manejar situaciones en las que el autor del archivo de formato de medios requiere que el reproductor o renderizador realice ciertas acciones al procesar el contenido de medios (por ejemplo, un flujo de bits de vídeo) asociado con el archivo de formato de medios. El esquema restringido permite a los reproductores inspeccionar el archivo de formato de medios para determinar fácilmente dichos requisitos para renderizar el contenido de medios y puede evitar que los reproductores heredados descarguen, descodifiquen y rendericen archivos que requieran procesamiento adicional. Un reproductor heredado puede incluir un dispositivo de reproducción que no está diseñado para procesar cierto contenido de medios porque, por ejemplo, el reproductor solo puede reproducir contenido de vídeo codificado sin el procesamiento requerido. Ejemplos ilustrativos de tipos de esquemas restringidos incluyen un esquema restringido para disposiciones de empaquetado de tramas de vídeo estereoscópico, un esquema restringido para vídeo de realidad virtual, entre otros.

[0005] Se imponen varias restricciones al esquema restringido del ISOBMFF. Por ejemplo, el número de esquemas restringidos que se pueden incluir en una parte de un archivo de formato de medios que es aplicable a cierto contenido de medios (por ejemplo, contenido de medios asociado con una entrada de muestras del archivo de formato de medios) está limitado a un solo esquema restringido. Tales restricciones no permiten que se apliquen simultáneamente varios tipos de esquemas restringidos al mismo contenido de medios asociado con una entrada de muestra. Una entrada de muestras se puede asociar con una o más muestras de un archivo de vídeo. Por ejemplo, cada flujo de medios puede estar contenido en una pista (del archivo de formato de medios) que está especializada para el tipo de medio del flujo de medios (por ejemplo, vídeo, audio u otro tipo de medios). El flujo de medios se puede parametrizar mediante una entrada de muestra, que contiene el nombre del tipo de medio (que indica el tipo de descodificador necesario para descodificar el flujo de medios) y cualquier parametrización del descodificador que sea necesaria. En un ejemplo ilustrativo, las restricciones sobre el esquema restringido en el ISOBMFF impiden la posibilidad de aplicar tanto el esquema restringido para disposiciones de vídeo estereoscópico como el esquema restringido para vídeo de realidad virtual al mismo contenido de vídeo asociado con una entrada de muestra. Sin embargo, el contenido de vídeo al que se aplican múltiples esquemas restringidos, como un vídeo de realidad virtual estereoscópico empaquetado en tramas, es posible y muy probable en aplicaciones de medios del mundo real.

65

[0006] Las técnicas y sistemas descritos en el presente documento pueden generar y procesar un archivo de formato de medios (de acuerdo con el formato de archivo de medios) que incluye múltiples esquemas restringidos que se pueden aplicar simultáneamente al mismo contenido de medios asociado con una entrada de muestra. Los términos esquema restringido y esquema de restricción pueden usarse indistintamente. El diseño mejorado se aplica a cualquier tipo de códec de vídeo. En algunos ejemplos, se pueden incluir cajas de tipos de esquema múltiples en una caja de información de esquema restringido contenida en una entrada de muestras de un archivo de formato de medios. Una caja de tipo de esquema define la naturaleza exacta de una restricción particular. Las cajas de tipos de esquema múltiples permiten la capacidad de identificar una pluralidad de esquemas restringidos requeridos para procesar el contenido de medios asociado con la entrada de muestra. Una caja de tipo de esquema también puede denominarse caja de tipo de esquema compatible. En algunos casos, también se pueden incluir cajas de información de esquemas múltiples en la caja de información de esquema restringido. Los datos necesarios para un esquema en particular se almacenan en la caja de información de esquema. Por ejemplo, se puede incluir una caja de información de esquema en la caja de información de esquema restringido para cada caja de tipo de esquema. En algunos ejemplos, se pueden incluir cajas de información de múltiples esquemas restringidos en un contenedor de entrada de muestra.

[0007] De acuerdo con al menos un ejemplo, se proporciona un procedimiento para procesar datos de vídeo. El procedimiento comprende obtener los datos de vídeo. El procedimiento comprende, además, generar un archivo de formato de medios que incluye los datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo.

[0008] En otro ejemplo, se proporciona un aparato para procesar datos de vídeo que incluye una memoria configurada para almacenar datos de vídeo y un procesador. El procesador está configurado y puede obtener los datos de vídeo. El procesador está configurado, además, para generar un archivo de formato de medios que incluye los datos de vídeo y una caja de información de esquema restringido asociada con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo.

[0009] En otro ejemplo, se proporciona un medio legible por ordenador no transitorio que tiene almacenadas instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que uno o más procesadores: obtengan los datos de vídeo; y generen un archivo de formato de medios que incluye los datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo, la caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema, en el que la pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquemas restringidos requeridos para procesar los datos de vídeo.

[0010] En otro ejemplo, se proporciona un aparato para procesar los datos de vídeo. El aparato incluye medios para obtener los datos de vídeo. El aparato incluye, además, medios para generar un archivo de formato de medios que incluye los datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo.

[0011] En algunos aspectos, cada caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica un tipo de esquema restringido diferente requerido para procesar los datos de vídeo.

[0012] En algunos aspectos, los datos de vídeo están asociados con una entrada de muestras del archivo de formato de medios.

[0013] En algunos aspectos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica una versión de un tipo de esquema restringido identificado por la caja de tipo de esquema.

[0014] En algunos aspectos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema incluye un identificador que identifica una fuente de datos necesarios para implementar un tipo de esquema restringido identificado por la caja de tipo de esquema.

[0015] En algunos aspectos, la caja de información de esquema restringido incluye, además, una pluralidad de cajas de información de esquema. Una caja de información de esquema de la pluralidad de cajas de información de esquema incluye los datos necesarios para un tipo de esquema restringido identificado por una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema.

[0016] En algunos aspectos, el archivo de formato de medios se genera usando un formato de archivo de medios de base ISO (ISOBMFF) o usando un formato de archivo de medios derivado del formato de archivo de medios de base ISO.

5 **[0017]** En algunos aspectos, el procedimiento, los aparatos y el medio legible por ordenador descritos anteriormente pueden comprender, además, transmitir el archivo de formato de medios.

[0018] En algunos aspectos, el procedimiento, los aparatos y el medio legible por ordenador descritos anteriormente pueden comprender, además, almacenar el archivo de formato de medios.

10 **[0019]** En algunos aspectos, el aparato comprende un dispositivo móvil con una cámara para capturar imágenes.

15 **[0020]** De acuerdo con al menos otro ejemplo, se proporciona un procedimiento de procesamiento de uno o más archivos de formato de medios. El procedimiento comprende obtener un archivo de formato de medios que incluye datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo. El procedimiento comprende, además, analizar el archivo de formato de medios y procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquemas restringidos identificados por la pluralidad de cajas de tipo de esquema.

20 **[0021]** En otro ejemplo, se proporciona un aparato para procesar uno o más archivos de formato de medios que incluye una memoria configurada para almacenar los datos de vídeo y un procesador. El procesador está configurado para, y puede obtener, un archivo de formato de medios que incluye datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo. El procesador está configurado, además, para poder analizar el archivo de formato de medios y procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquema restringido identificados por la pluralidad de cajas de tipo de esquema.

25 **[0022]** En otro ejemplo, se proporciona un medio legible por ordenador no transitorio que tiene almacenadas instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que uno o más procesadores: obtengan un archivo de formato de medios que incluya datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo, incluyendo la caja de información de esquema restringido una pluralidad de cajas de tipo de esquema, en el que la pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo; analizar el archivo de formato de medios; y procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquema restringido identificados por la pluralidad de cajas de tipo de esquema.

30 **[0023]** En otro ejemplo, se proporciona un aparato para procesar uno o más archivos de formato de medios. El aparato incluye medios para obtener un archivo de formato de medios que incluye datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo. El aparato incluye, además, medios para analizar el archivo de formato de medios y medios para procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquema restringido identificados por la pluralidad de cajas de tipo de esquema.

35 **[0024]** En algunos aspectos, cada caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica un tipo de esquema restringido diferente requerido para procesar los datos de vídeo.

40 **[0025]** En algunos aspectos, los datos de vídeo están asociados con una entrada de muestras del archivo de formato de medios.

[0026] En algunos aspectos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica una versión de un tipo de esquema restringido identificado por la caja de tipo de esquema.

45 **[0027]** En algunos aspectos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema incluye un identificador que identifica una fuente de datos necesarios para implementar un tipo de esquema restringido identificado por la caja de tipo de esquema.

50 **[0028]** En algunos aspectos, la caja de información de esquema restringido incluye, además, una pluralidad de cajas de información de esquema. Una caja de información de esquema de la pluralidad de cajas de información

de esquema incluye los datos necesarios para un tipo de esquema restringido identificado por una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema.

[0029] En algunos aspectos, el archivo de formato de medios se genera usando un formato de archivo de medios de base ISO (ISOBMFF) o usando un formato de archivo de medios derivado del formato de archivo de medios de base ISO.

[0030] En algunos aspectos, el procedimiento, los aparatos y el medio legible por ordenador descritos anteriormente pueden comprender, además, descodificar los datos de vídeo. Por ejemplo, el aparato puede incluir un descodificador para descodificar los datos de vídeo.

[0031] En algunos aspectos, el procedimiento, los aparatos y el medio legible por ordenador descritos anteriormente pueden comprender, además, mostrar los datos de vídeo descodificados. Por ejemplo, el aparato puede incluir una pantalla para visualizar los datos de vídeo descodificados.

[0032] En algunos aspectos, el aparato comprende un dispositivo móvil con una cámara para capturar imágenes.

[0033] Esta breve explicación no pretende identificar rasgos característicos clave o esenciales de la materia objeto reivindicada, ni pretende usarse de forma aislada para determinar el alcance de la materia objeto reivindicada. La materia objeto debe entenderse por referencia a las partes apropiadas de toda la memoria descriptiva de esta patente, cualquiera o todos los dibujos, y cada reivindicación.

[0034] Lo anterior, junto con otros rasgos característicos y modos de realización, resultará más evidente tras la referencia a la siguiente memoria descriptiva, reivindicaciones y dibujos adjuntos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0035] A continuación, se describen en detalle modos de realización ilustrativos de la presente invención con referencia a las siguientes figuras de dibujos:

la FIG. 1 es un diagrama de bloques que ilustra un ejemplo de un dispositivo de codificación y un dispositivo de descodificación, de acuerdo con algunos ejemplos;

la FIG. 2 es un diagrama que ilustra una estructura de archivo de ejemplo que sigue el formato de archivo de medios de base ISO (ISOBMFF), de acuerdo con algunos ejemplos;

la FIG. 3 es un diagrama que ilustra un ejemplo de un archivo de medios de base ISO (formateado de acuerdo con ISOBMFF) que contiene datos y metadatos para una presentación de vídeo, de acuerdo con algunos ejemplos;

la FIG. 4 es un diagrama que ilustra un ejemplo de una caja de medios que se puede incluir en un archivo de medios de base ISO, de acuerdo con algunos ejemplos;

la FIG. 5 es un diagrama de flujo que ilustra un ejemplo de un proceso para procesar datos de vídeo, de acuerdo con algunos ejemplos;

la FIG. 6 es un diagrama de flujo que ilustra un ejemplo de un proceso para procesar uno o más archivos de formato de medios, de acuerdo con algunos ejemplos;

la FIG. 7 es un diagrama de bloques que ilustra un dispositivo de codificación de vídeo de ejemplo, de acuerdo con algunos ejemplos; y

la FIG. 8 es un diagrama de bloques que ilustra un dispositivo de descodificación de vídeo de ejemplo, de acuerdo con algunos ejemplos.

DESCRIPCIÓN DETALLADA

[0036] A continuación, se proporcionan determinados aspectos y modos de realización de esta divulgación. Algunos de estos aspectos y modos de realización pueden aplicarse de manera independiente y algunos de ellos se pueden aplicar en combinación, como reconocerá un experto en la técnica. En la siguiente descripción se exponen, con fines explicativos, detalles específicos para permitir un entendimiento exhaustivo de los modos de realización de la invención. Sin embargo, resultará evidente que diversos modos de realización pueden llevarse a la práctica sin estos detalles específicos. Las figuras y la descripción no pretenden ser restrictivas.

[0037] La siguiente descripción solo proporciona modos de realización a modo de ejemplo y no pretende limitar el alcance, la aplicabilidad o la configuración de la divulgación. En cambio, la siguiente descripción de los modos de realización a modo de ejemplo proporcionará a los expertos en la técnica una descripción que permite implementar un modo de realización a modo de ejemplo. Debe entenderse que se pueden hacer diversos cambios en la función y disposición de los elementos sin apartarse del alcance de la invención como se establece en las reivindicaciones adjuntas.

[0038] En la siguiente descripción se ofrecen detalles específicos para proporcionar un entendimiento exhaustivo de los modos de realización. Sin embargo, un experto en la técnica entenderá que los modos de realización se pueden llevar a la práctica sin estos detalles específicos. Por ejemplo, circuitos, sistemas, redes, procesos y otros componentes pueden mostrarse como componentes en forma de diagrama de bloques para no oscurecer los modos de realización con detalles innecesarios. En otros casos, circuitos, procesos, algoritmos, estructuras y técnicas ampliamente conocidos se pueden mostrar sin detalles innecesarios para no oscurecer los modos de realización.

[0039] Además, cabe destacar que los modos de realización individuales se pueden describir como un proceso que se representa como un organigrama, un diagrama de flujo, un diagrama de flujo de datos, un diagrama de estructura o un diagrama de bloques. Aunque un organigrama puede describir las operaciones como un proceso secuencial, muchas de las operaciones pueden realizarse en paralelo o simultáneamente. Además, el orden de las operaciones puede disponerse. Un proceso se termina cuando se completan sus operaciones, pero podría tener etapas adicionales no incluidas en una figura. Un proceso puede corresponder a un procedimiento, una función, un procedimiento, una subrutina, un subprograma, etc. Cuando un proceso corresponde a una función, su terminación puede corresponder con el retorno de la función a la función de llamada o a la función principal.

[0040] La expresión "medio legible por ordenador" incluye, pero no se limita a, dispositivos de almacenamiento portátiles o no portátiles, dispositivos de almacenamiento óptico y otros diversos medios que puedan almacenar, contener o transportar instrucciones y/o datos. Un medio legible por ordenador puede incluir un medio no transitorio en el que se pueden almacenar datos y que no incluye ondas portadoras y/o señales electrónicas transitorias que se propagan de forma inalámbrica o por conexiones por cable. Ejemplos de un medio no transitorio pueden incluir, pero no se limitan a, un disco o cinta magnéticos, medios de almacenamiento óptico tales como un disco compacto (CD) o un disco versátil digital (DVD), memoria flash, memoria o dispositivos de memoria. Un medio legible por ordenador puede tener almacenado en el mismo código y/o instrucciones ejecutables por máquina que pueden representar un procedimiento, una función, un subprograma, un programa, una rutina, una subrutina, un módulo, un paquete de software, una clase o cualquier combinación de instrucciones, estructuras de datos o sentencias de programa. Un segmento de código se puede acoplar a otro segmento de código o a un circuito de hardware pasando y/o recibiendo información, datos, argumentos, parámetros o contenido de memoria. Se pueden pasar, reenviar o transmitir información, argumentos, parámetros, datos, etc. a través de cualquier medio adecuado, incluyendo compartición de memoria, paso de mensajes, paso de testigos, transmisión por red o similares.

[0041] Además, los modos de realización se pueden implementar mediante hardware, software, firmware, middleware, microcódigo, lenguajes de descripción de hardware o cualquier combinación de los mismos. Cuando se implementan en software, firmware, middleware o microcódigo, el código de programa o los segmentos de código para realizar las tareas necesarias (por ejemplo, un producto de programa informático) pueden almacenarse en un medio legible por ordenador o legible por máquina. Un(os) procesador(es) puede(n) realizar las tareas necesarias.

[0042] El contenido de vídeo se puede capturar y codificar en un flujo de bits de vídeo codificado. Los formatos de medios se pueden utilizar para encapsular vídeo codificado. Un ejemplo de formato de medios incluye el formato de archivo de medios de base ISO (ISOBMFF). El ISOBMFF se usa como base para muchos formatos de encapsulación de códecs. Los medios continuos (por ejemplo, audio y vídeo), los medios estáticos (por ejemplo, imágenes), así como los metadatos se pueden almacenar en un archivo que cumpla con ISOBMFF. Los archivos estructurados de acuerdo con el ISOBMFF se pueden usar para muchos propósitos, incluyendo la reproducción local de archivos de medios, la descarga progresiva de un archivo remoto, segmentos para la transmisión en continuo dinámica adaptativa a través de HTTP (DASH), contenedores para contenido que se va a transmitir en continuo y las instrucciones de empaquetado para el contenido y el registro de flujos de medios recibidos en tiempo real, entre otros.

[0043] El ISOBMFF incluye un diseño de esquema restringido que se usa cuando un creador (por ejemplo, un autor, productor o similar) del archivo de formato de medios requiere que el reproductor o renderizador realice ciertas acciones al procesar contenido de medios (por ejemplo, flujo de bits de vídeo codificado) asociado con el archivo de formato de medios. Como se describe con más detalle a continuación, uno o más sistemas y procedimientos descritos en el presente documento están dirigidos a mejorar el diseño del esquema restringido. Estas mejoras se pueden usar para archivos basados en ISOBMFF o para cualquier otro formato de archivo.

[0044] Como se señaló anteriormente, las tramas de vídeo capturadas se pueden codificar y/o comprimir para almacenamiento y/o transmisión, y un dispositivo receptor puede decodificar y/o descomprimir las tramas de vídeo codificadas. A medida que más dispositivos y sistemas ofrecen a los consumidores la capacidad de consumir datos de vídeo digital, la necesidad de técnicas de codificación de vídeo eficientes se vuelve más importante. La codificación de vídeo es necesaria para reducir los requisitos de almacenamiento y transmisión necesarios para manejar las grandes cantidades de datos presentes en los datos de vídeo digital. Se pueden usar diversas técnicas de codificación de vídeo para comprimir los datos de vídeo en una forma que use una velocidad de transmisión de bits más baja mientras mantiene una alta calidad de vídeo.

[0045] La FIG. 1 es un diagrama de bloques que ilustra un ejemplo de un sistema 100 de codificación de vídeo que incluye un dispositivo de codificación 104 y un dispositivo de decodificación 112. El dispositivo de codificación 104 puede ser parte de un dispositivo de origen, y el dispositivo de decodificación 112 puede ser parte de un dispositivo de recepción. El dispositivo de origen y/o el dispositivo de recepción pueden incluir un dispositivo electrónico, como un teléfono móvil o estacionario (por ejemplo, teléfono inteligente, teléfono celular o similar), un ordenador de escritorio, un ordenador portátil o notebook, una tableta electrónica, un decodificador, un televisor, una cámara, un dispositivo de visualización, un reproductor de medios digitales, una consola de videojuegos, un dispositivo de transmisión de vídeo en tiempo real, una cámara de protocolo de Internet (IP) o cualquier otro dispositivo electrónico adecuado. En algunos ejemplos, el dispositivo de origen y el dispositivo de recepción pueden incluir uno o más transceptores inalámbricos para comunicaciones inalámbricas. Las técnicas de codificación descritas en el presente documento son aplicables a la codificación de vídeo en varias aplicaciones multimedia, incluidas las transmisiones continuas de vídeo (por ejemplo, a través de Internet), radiodifusiones o transmisiones de televisión, codificación de vídeo digital para almacenamiento en un medio de almacenamiento de datos, decodificación de vídeo digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema 100 puede admitir una transmisión de vídeo unidireccional o bidireccional para admitir aplicaciones tales como videoconferencia, la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0046] El dispositivo de codificación 104 (o codificador) se puede usar para codificar datos de vídeo usando una norma o un protocolo de codificación de vídeo para generar un flujo de bits de vídeo codificado. Ejemplos de las normas de codificación de vídeo incluyen ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluyendo sus ampliaciones de Codificación de Vídeo Escalable (SVC) y de Codificación de Vídeo Multivista (MVC) y la Codificación de Vídeo de Alta Eficiencia (HEVC) o ITU-T H.265. Existen varias ampliaciones para HEVC que tratan con codificación de vídeo multicapa, incluidas las ampliaciones de codificación de contenido de pantalla y rango, codificación de vídeo 3D (3D-HEVC) y ampliaciones multivista (MV-HEVC) y ampliación escalable (SHVC). El HEVC y sus ampliaciones han sido desarrollados por el Equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC), así como el Equipo de colaboración conjunta sobre el desarrollo de ampliaciones de codificación de vídeo 3D (JCT-3V) del Grupo de expertos en codificación de vídeo del ITU-T (VCEG) y el Grupo de expertos en imágenes en movimiento ISO/IEC (MPEG). MPEG e ITU-T VCEG también han formado un equipo de vídeo de exploración conjunta (JVET) para explorar nuevas herramientas de codificación para la próxima generación de estándares de codificación de vídeo. El software de referencia se llama JEM (modelo de exploración conjunta).

[0047] Muchos modos de realización descritos en el presente documento proporcionan ejemplos usando el modelo JEM, la norma HEVC y/o ampliaciones de los mismos. Sin embargo, las técnicas y sistemas descritos en el presente documento también pueden ser aplicables a otros estándares de codificación, como AVC, MPEG, ampliaciones de los mismos u otros estándares de codificación adecuados ya disponibles o aún no disponibles o desarrollados. En consecuencia, aunque las técnicas y sistemas descritos en el presente documento pueden describirse con referencia a una norma de codificación de vídeo particular, un experto en la técnica apreciará que la descripción no debe interpretarse para que se aplique solo a esa norma particular.

[0048] Haciendo referencia a la FIG. 1, una fuente de vídeo 102 puede proporcionar los datos de vídeo al dispositivo de codificación 104. La fuente de vídeo 102 puede ser parte del dispositivo de origen, o puede ser parte de un dispositivo que no sea el dispositivo de origen. La fuente de vídeo 102 puede incluir un dispositivo de captura de vídeo (por ejemplo, una videocámara, un teléfono con cámara, un videoteléfono o similar), un archivo de vídeo que contiene vídeo almacenado, un servidor de vídeo o proveedor de contenido que proporciona datos de vídeo, una interfaz de alimentación de vídeo que recibe vídeo desde un servidor de vídeo o proveedor de contenido, un sistema de gráficos de ordenador para generar datos de vídeo de gráficos de ordenador, una combinación de dichas fuentes o cualquier otra fuente de vídeo adecuada.

[0049] Los datos de vídeo de la fuente de vídeo 102 pueden incluir una o más tramas o imágenes de entrada. Cada imagen o trama de un vídeo es una imagen fija de una escena. En algunos ejemplos, una imagen o trama de los datos de vídeo puede incluir datos de vídeo de 360 grados correlacionados o proyectados en planos de una geometría (por ejemplo, un cubo u otra geometría adecuada). El motor codificador 106 (o codificador) del dispositivo de codificación 104 codifica los datos de vídeo para generar un flujo de bits de vídeo codificado. En algunos ejemplos, un flujo de bits de vídeo codificado (o "flujo de bits de vídeo" o "flujo de bits") es

una serie de una o más secuencias de vídeo codificadas. Una secuencia de vídeo codificada (CVS) incluye una serie de unidades de acceso (AU) que comienzan con una AU que tiene una imagen de punto de acceso aleatorio en la capa base y con ciertas propiedades hasta y sin incluir una siguiente AU que tenga una imagen de punto de acceso aleatorio en la capa base y con ciertas propiedades. Por ejemplo, entre las ciertas propiedades de una imagen de punto de acceso aleatorio que inicia una CVS puede incluirse un indicador RASL (por ejemplo, NoRaslOutputFlag) igual a 1. De lo contrario, una imagen de punto de acceso aleatorio (con el indicador RASL igual a 0) no inicia una CVS. Una unidad de acceso (AU) incluye una o más imágenes codificadas e información de control correspondiente a las imágenes codificadas que comparten el mismo tiempo de salida. Los fragmentos de imágenes codificados se encapsulan en el nivel de flujo de bits en unidades de datos llamadas unidades de capa de abstracción de red (NAL). Por ejemplo, un flujo de bits de vídeo HEVC puede incluir una o más CVS incluyendo unidades NAL. Cada una de las unidades NAL tiene una cabecera de unidad NAL. En un ejemplo, la cabecera es de un byte para H.264/AVC (excepto para ampliaciones de varias capas) y dos bytes para HEVC. Los elementos de sintaxis en la cabecera de la unidad NAL toman los bits designados y, por lo tanto, son visibles para todo tipo de sistemas y capas de transporte, como flujo de transporte, protocolo de transporte en tiempo real (RTP), formato de archivo, entre otros.

[0050] Existen dos clases de unidades NAL en la norma HEVC, incluyendo unidades NAL de capa de codificación de vídeo (VCL) y unidades NAL no VCL. Una unidad NAL VCL incluye un fragmento o segmento de fragmento (descrito a continuación) de datos de imagen codificados, y una unidad NAL no VCL incluye información de control que se relaciona con una o más imágenes codificadas. En algunos casos, una unidad NAL puede denominarse paquete. Una AU HEVC incluye unidades NAL VCL que contienen datos de imagen codificados y unidades NAL no VCL (si las hay) correspondientes a los datos de imagen codificados.

[0051] Las unidades NAL pueden contener una secuencia de bits que forman una representación codificada de los datos de vídeo (por ejemplo, un flujo de bits de vídeo codificado, una CVS de un flujo de bits o similar), como representaciones codificadas de imágenes en un vídeo. El motor codificador 106 genera representaciones codificadas de imágenes al dividir cada imagen en múltiples fragmentos. Un fragmento es independiente de otros fragmentos, de modo que la información en el fragmento se codifica independientemente de los datos de otros fragmentos dentro de la misma imagen. Un fragmento incluye uno o más segmentos de fragmento que incluyen un segmento de fragmento independiente y, si están presentes, uno o más segmentos de fragmento dependientes que dependen de segmentos de fragmento anteriores. Los fragmentos se dividen en bloques de árbol de codificación (CTB) de muestras de luma y muestras de croma. Un CTB de muestras de luma y uno o más CTB de muestras de croma, junto con la sintaxis de las muestras, se denominan unidad de árbol de codificación (CTU). Una CTU es la unidad de procesamiento básica para la codificación HEVC. Una CTU se puede dividir en múltiples unidades de codificación (CU) de diferentes tamaños. Una CU contiene matrices de muestras de luma y croma que se denominan bloques de codificación (CB).

[0052] Los CB de luma y croma pueden dividirse adicionalmente en bloques de predicción (PB). Un PB es un bloque de muestras del componente luma o un componente croma que utiliza los mismos parámetros de movimiento para la interpredicción o predicción de la copia intrabloque (cuando está disponible o habilitado para su uso). El PB de luma y uno o más PB de croma, junto con la sintaxis asociada, forman una unidad de predicción (PU). Para la predicción inter, un conjunto de parámetros de movimiento (por ejemplo, uno o más vectores de movimiento, índices de referencia o similares) se señalan en el flujo de bits para cada PU y se usan para la interpredicción de PB de luma y el uno o más PB de croma. Los parámetros de movimiento también se pueden denominar información de movimiento. Un CB también se puede dividir en uno o más bloques de transformación (TB). Un TB representa un bloque cuadrado de muestras de una componente de color en la que se aplica la misma transformación bidimensional para codificar una señal residual de predicción. Una unidad de transformación (TU) representa los TB de las muestras de luma y croma, así como elementos de sintaxis correspondientes.

[0053] Un tamaño de la CU corresponde a un tamaño del modo de codificación y puede ser de forma cuadrada. Por ejemplo, el tamaño de una CU puede ser de 8 x 8 muestras, 16 x 16 muestras, 32 x 32 muestras, 64 x 64 muestras, o cualquier otro tamaño apropiado hasta el tamaño de la CTU correspondiente. La expresión "N x N" se usa en el presente documento para referirse a las dimensiones en píxeles de un bloque de vídeo en términos de dimensiones verticales y horizontales (por ejemplo, 8 píxeles x 8 píxeles). Los píxeles de un bloque se pueden disponer en filas y columnas. En algunos modos de realización, no es necesario que los bloques tengan el mismo número de píxeles en una dirección horizontal y en una dirección vertical. Los datos de sintaxis asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir en si la CU está codificada en modo de intrapredicción o codificada en modo de interpredicción. Las PU se pueden dividir para tener una conformación no cuadrada. Los datos de sintaxis asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con una CTU. Una TU puede tener una conformación cuadrada o no cuadrada.

[0054] De acuerdo con la norma HEVC, las transformaciones se pueden realizar usando unidades de transformación (TU). Las TU pueden variar para diferentes CU. Las TU pueden dimensionarse en función del tamaño de las PU dentro de una CU dada. Las TU pueden ser del mismo tamaño o de un tamaño más pequeño

que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU se pueden subdividir en unidades más pequeñas usando una estructura de árbol cuaternario conocida como árbol cuaternario residual (RQT). Los nodos hoja del RQT pueden corresponder a las TU. Los valores de diferencia de píxeles asociados a las TU se pueden transformar para producir coeficientes de transformación. A continuación, los coeficientes de transformación pueden ser cuantificados por el motor codificador 106.

[0055] Una vez que las imágenes de los datos de vídeo se dividen en CU, el motor codificador 106 predice cada PU usando un modo de predicción. A continuación, la unidad de predicción o el bloque de predicción se resta de los datos de vídeo originales para obtener residuos (descritos a continuación). Para cada CU, se puede señalar un modo de predicción dentro del flujo de bits usando datos de sintaxis. Un modo de predicción puede incluir intrapredicción (o predicción intraimagen) o interpredicción (o predicción interimagen). La intrapredicción utiliza la correlación entre las muestras espacialmente contiguas dentro de una imagen. Por ejemplo, usando la intrapredicción, cada PU se predice a partir de datos de imágenes contiguas en la misma imagen usando, por ejemplo, predicción de DC para encontrar un valor promedio para la PU, predicción plana para ajustar una superficie plana a la PU, predicción de dirección para extrapolar a partir de los datos contiguos, o cualquier otro tipo adecuado de predicción. La interpredicción usa la correlación temporal entre imágenes para obtener una predicción de compensación de movimiento para un bloque de muestras de imágenes. Por ejemplo, usando la interpredicción, cada PU se predice usando predicción de compensación de movimiento a partir de los datos de imagen en una o más imágenes de referencia (antes o después de la imagen actual en orden de salida). La decisión de codificar un área de imagen usando predicción interimagen o intraimagen puede tomarse, por ejemplo, a nivel de CU.

[0056] En algunos ejemplos, al uno o más fragmentos de una imagen se le asigna un tipo de fragmento. Entre los tipos de fragmentos se incluye un fragmento I, un fragmento P y un fragmento B. Un fragmento I (intra-tramas, decodificable independientemente) es un fragmento de una imagen que solo está codificado por intrapredicción, y por lo tanto es decodificable de forma independiente ya que el fragmento I requiere solo los datos dentro de la trama para predecir cualquier unidad de predicción o bloque de predicción del fragmento. Un fragmento P (tramas predichas unidireccionales) es un fragmento de una imagen que puede codificarse con intrapredicción y con interpredicción unidireccionales. Cada unidad de predicción o bloque de predicción dentro de un fragmento P está codificado con intrapredicción o interpredicción. Cuando se aplica la interpredicción, la unidad de predicción o el bloque de predicción solo se predicen mediante una imagen de referencia y, por lo tanto, las muestras de referencia son solo de una región de referencia de una trama. Un fragmento B (tramas predictivas bidireccionales) es un fragmento de una imagen que puede codificarse con intrapredicción e interpredicción (por ejemplo, bipredicción o unipredicción). Una unidad de predicción o un bloque de predicción de un fragmento B puede predecirse bidireccionalmente a partir de dos imágenes de referencia, donde cada imagen contribuye con una región de referencia y los conjuntos de muestras de las dos regiones de referencia se ponderan (por ejemplo, con ponderaciones iguales o con ponderaciones diferentes) para producir la señal de predicción del bloque predicho bidireccional. Como se expuso anteriormente, los fragmentos de una imagen se codifican de forma independiente. En algunos casos, una imagen puede codificarse como un solo fragmento.

[0057] Una PU puede incluir los datos (por ejemplo, parámetros de movimiento u otros datos adecuados) relacionados con el proceso de predicción. Por ejemplo, cuando la PU se codifica usando intrapredicción, la PU puede incluir datos que describen un modo de intrapredicción para la PU. Como otro ejemplo, cuando la PU se codifica usando interpredicción, la PU puede incluir datos que definen un vector de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, una componente horizontal del vector de movimiento (Δx), una componente vertical del vector de movimiento (Δy), una resolución para el vector de movimiento (por ejemplo, precisión entera, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento, un índice de referencia, una lista de imágenes de referencia (por ejemplo, lista 0, lista 1 o lista C) para el vector de movimiento, o cualquier combinación de los mismos.

[0058] El dispositivo de codificación 104 puede entonces realizar la transformación y cuantificación. Por ejemplo, tras la predicción, el motor codificador 106 puede calcular valores residuales correspondientes a la PU. Los valores residuales pueden comprender valores de diferencia de píxeles entre el bloque actual de píxeles que se codifican (la PU) y el bloque de predicción usado para predecir el bloque actual (por ejemplo, la versión predicha del bloque actual). Por ejemplo, después de generar un bloque de predicción (por ejemplo, emitir interpredicción o intrapredicción), el motor codificador 106 puede generar un bloque residual restando el bloque de predicción producido por una unidad de predicción del bloque actual. El bloque residual incluye un conjunto de valores de diferencia de píxeles que cuantifican las diferencias entre los valores de píxeles del bloque actual y los valores de píxeles del bloque de predicción. En algunos ejemplos, el bloque residual puede representarse en un formato de bloque bidimensional (por ejemplo, una matriz bidimensional o una matriz de valores de píxeles). En tales ejemplos, el bloque residual es una representación bidimensional de los valores de píxel.

[0059] Todos los datos residuales que pueden quedar después de que se lleve a cabo la predicción se transforman usando una transformación de bloque, que puede basarse en transformación cosenoidal discreta, transformación sinusoidal discreta, una transformación entera, una transformación de tren de ondas, otra función

de transformación adecuada, o cualquier combinación de las mismas. En algunos casos, una o más transformaciones de bloque (por ejemplo, tamaños 32 x 32, 16 x 16, 8 x 8, 4 x 4 u otro tamaño adecuado) pueden aplicarse a los datos residuales en cada CU. En algunos modos de realización, se puede usar una TU para los procesos de transformación y cuantificación implementados por el motor codificador 106. Una CU dada que tiene una o más PU también puede incluir una o más TU. Como se describe con más detalle a continuación, los valores residuales pueden transformarse en coeficientes de transformación usando las transformaciones de bloque, y después pueden cuantificarse y escanearse usando TU para producir coeficientes de transformación serializados para la codificación por entropía.

[0060] En algunos modos de realización, tras la codificación de intrapredicción o interpredicción mediante las PU de una CU, el motor codificador 106 puede calcular los datos residuales para las TU de la CU. Las PU pueden comprender datos de píxel en el dominio espacial (o dominio de píxel). Las TU pueden comprender coeficientes en el dominio de transformación después de la aplicación de una transformación de bloque. Como se ha observado previamente, los datos residuales pueden corresponder a valores de diferencia de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las PU. El motor codificador 106 puede formar las TU, incluyendo los datos residuales para la CU y, a continuación, transformar las TU para producir coeficientes de transformación para la CU.

[0061] El motor codificador 106 puede realizar la cuantificación de los coeficientes de transformación. La cuantificación proporciona una compresión adicional al cuantificar los coeficientes de transformación para reducir la cantidad de datos usados para representar los coeficientes. Por ejemplo, la cuantificación puede reducir la profundidad de bits asociada a algunos o todos los coeficientes. En un ejemplo, un coeficiente con un valor de n bits puede redondearse a la baja hasta un valor de m bits durante la cuantificación, donde n es mayor que m.

[0062] Una vez que se lleva a cabo la cuantificación, el flujo de bits de vídeo codificado incluye coeficientes de transformación cuantificados, información de predicción (por ejemplo, modos de predicción, vectores de movimiento, vectores de bloques, o similares), información de división, y cualesquiera otros datos adecuados, tales como otros datos de sintaxis. Los diferentes elementos del flujo de bits de vídeo codificado pueden ser codificados por entropía por el motor codificador 106. En algunos ejemplos, el motor codificador 106 puede utilizar un orden de exploración predefinido para explorar los coeficientes de transformación cuantificados para producir un vector serializado que se pueda codificar por entropía. En algunos ejemplos, el motor codificador 106 puede realizar una exploración adaptativa. Después de escanear los coeficientes de transformación cuantificados para formar un vector (por ejemplo, un vector unidimensional), el motor codificador 106 puede codificar por entropía el vector. Por ejemplo, el motor de codificador 106 puede usar codificación de longitud variable adaptativa al contexto, codificación aritmética binaria adaptativa al contexto, codificación aritmética binaria adaptativa al contexto basada en sintaxis, codificación por entropía por división en intervalos de probabilidad u otra técnica de codificación por entropía adecuada.

[0063] Como se ha descrito previamente, un flujo de bits HEVC incluye un grupo de unidades NAL, incluyendo unidades NAL VCL y unidades NAL no VCL. Las unidades NAL VCL incluyen datos de imagen codificados que forman un flujo de bits de vídeo codificado. Por ejemplo, una secuencia de bits que forman el flujo de bits de vídeo codificado está presente en las unidades NAL VCL. Las unidades NAL no VCL pueden contener conjuntos de parámetros con información de alto nivel relacionada con el flujo de bits de vídeo codificado, además de otra información. Por ejemplo, un conjunto de parámetros puede incluir un conjunto de parámetros de vídeo (VPS), un conjunto de parámetros de secuencia (SPS) y un conjunto de parámetros de imagen (PPS). Ejemplos de objetivos de los conjuntos de parámetros incluyen la eficiencia de la velocidad de transferencia de bits, la resistencia a los errores y el suministro de interfaces de capa de sistemas. Cada fragmento hace referencia a un único PPS, SPS y VPS activo para acceder a la información que el dispositivo de descodificación 112 puede usar para descodificar el fragmento. Se puede codificar un identificador (ID) para cada conjunto de parámetros, incluyendo un ID de VPS, un ID de SPS y un ID de PPS. Un SPS incluye un ID de SPS y un ID de VPS. Un PPS incluye un ID de PPS y un ID de SPS. Cada cabecera de fragmento incluye un ID de PPS. Usando los ID, se pueden identificar conjuntos de parámetros activos para un fragmento dado.

[0064] Un PPS incluye información que se aplica a todos los fragmentos en una imagen dada. Debido a esto, todos los fragmentos en una imagen se refieren al mismo PPS. Los fragmentos en diferentes imágenes también pueden referirse al mismo PPS. Un SPS incluye información que se aplica a todas las imágenes en una misma secuencia de vídeo codificada (CVS) o flujo de bits. Como se describió previamente, una secuencia de vídeo codificada es una serie de unidades de acceso (AU) que comienza con una imagen de punto de acceso aleatorio (por ejemplo, una imagen de referencia de descodificación instantánea (IDR) o una imagen de acceso de enlace roto (BLA) u otra imagen de punto de acceso aleatorio apropiada) en la capa base y con ciertas propiedades (descritas anteriormente) hasta y sin incluir una siguiente AU que tenga una imagen de punto de acceso aleatorio en la capa base y con ciertas propiedades (o el final del flujo de bits). La información en un SPS puede no cambiar de una imagen a otra dentro de una secuencia de vídeo codificada. Las imágenes en una secuencia de vídeo codificada pueden usar el mismo SPS. El VPS incluye información que se aplica a todas las capas dentro de una secuencia de vídeo codificada o flujo de bits codificado. El VPS incluye una estructura de sintaxis con elementos de sintaxis que se aplican a secuencias de vídeo codificadas completas. En algunos modos de

realización, el VPS, el SPS o el PPS pueden transmitirse en banda con el flujo de bits codificado. En algunos modos de realización, el VPS, el SPS o el PPS pueden transmitirse fuera de banda en una transmisión distinta de la de las unidades NAL que contienen datos de vídeo codificados.

[0065] Un flujo de bits de vídeo también puede incluir mensajes de información complementaria de mejora (SEI). Por ejemplo, una unidad NAL SEI puede ser parte del flujo de bits de vídeo. En algunos casos, un mensaje SEI puede contener información que no es necesaria para el proceso de descodificación. Por ejemplo, la información en un mensaje SEI puede no ser esencial para que el descodificador descodifique las imágenes de vídeo del flujo de bits, pero el descodificador puede usar la información para mejorar la visualización o el procesamiento de las imágenes (por ejemplo, la salida descodificada). La información de un mensaje SEI puede ser metadatos incrustados. En un ejemplo ilustrativo, las entidades del lado del descodificador podrían usar la información de un mensaje SEI para mejorar la visibilidad del contenido. En algunos casos, ciertos estándares de aplicación pueden exigir la presencia de dichos mensajes SEI en el flujo de bits para que la mejora de la calidad se pueda llevar a todos los dispositivos que cumplen con el estándar de aplicación (por ejemplo, el transporte del mensaje SEI de empaquetado en tramas para formato de vídeo 3DTV plano-estereoscópico compatible, donde el mensaje SEI se transmite para cada trama del vídeo, manejo de un mensaje SEI del punto de recuperación, uso de mensaje SEI de rectángulo de escaneo panorámico en DVB, además de muchos otros ejemplos).

[0066] La salida 110 del dispositivo de codificación 104 puede enviar las unidades NAL que componen los datos del flujo de bits de vídeo codificados a través del enlace de comunicaciones 120 al dispositivo de descodificación 112 del dispositivo de recepción. La entrada 114 del dispositivo de descodificación 112 puede recibir las unidades NAL. El enlace de comunicaciones 120 puede incluir un canal proporcionado por una red inalámbrica, una red por cable o una combinación de una red por cable e inalámbrica. Una red inalámbrica puede incluir cualquier interfaz inalámbrica o combinación de interfaces inalámbricas y puede incluir cualquier red inalámbrica adecuada (por ejemplo, Internet u otra red de área amplia, una red basada en paquetes, WiFi™, radiofrecuencia (RF), UWB, WiFi-Direct, celular, Evolución a Largo Plazo (LTE), WiMax™ o similar). Una red por cable puede incluir cualquier interfaz por cable (por ejemplo, fibra, ethernet, ethernet de línea eléctrica, ethernet sobre cable coaxial, línea de señal digital (DSL) o similar). Las redes por cable y/o inalámbricas pueden implementarse usando diversos equipos, tales como estaciones base, encaminadores, puntos de acceso, puentes, pasarelas, conmutadores o similares. Los datos de flujo de bits de vídeo codificado se pueden modular de acuerdo con un estándar de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo de recepción.

[0067] En algunos ejemplos, el dispositivo de codificación 104 puede almacenar datos de flujo de bits de vídeo codificados en el almacenamiento 108. La salida 110 puede recuperar los datos de flujo de bits de vídeo codificado del motor codificador 106 o del almacenamiento 108. El almacenamiento 108 puede incluir cualquiera de una variedad de medios de almacenamiento de datos distribuidos o de acceso local. Por ejemplo, el almacenamiento 108 puede incluir un disco duro, un disco de almacenamiento, memoria flash, memoria volátil o no volátil o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de vídeo codificados.

[0068] La entrada 114 del dispositivo de descodificación 112 recibe los datos de flujo de bits de vídeo codificado y puede proporcionar los datos de flujo de bits de vídeo al motor descodificador 116, o al almacenamiento 118 para su uso posterior por parte del motor descodificador 116. El motor descodificador 116 puede descodificar los datos de flujo de bits de vídeo codificado mediante descodificación por entropía (por ejemplo, usando un descodificador por entropía) y extrayendo los elementos de una o más secuencias de vídeo codificadas que forman los datos de vídeo codificados. A continuación, el motor descodificador 116 puede reescalar y realizar una transformación inversa en los datos de flujo de bits de vídeo codificado. A continuación, los datos residuales se pasan a una etapa de predicción del motor descodificador 116. A continuación, el motor descodificador 116 predice un bloque de píxeles (por ejemplo, una PU). En algunos ejemplos, la predicción se agrega a la salida de la transformación inversa (los datos residuales).

[0069] El dispositivo de descodificación 112 puede emitir el vídeo descodificado a un dispositivo de destino de vídeo 122, que puede incluir una pantalla u otro dispositivo de salida para mostrar los datos de vídeo descodificados a un consumidor del contenido. En algunos aspectos, el dispositivo de destino de vídeo 122 puede ser parte del dispositivo de recepción que incluye el dispositivo de descodificación 112. En algunos aspectos, el dispositivo de destino de vídeo 122 puede ser parte de un dispositivo separado que no sea el dispositivo de recepción.

[0070] En algunos modos de realización, el dispositivo de codificación de vídeo 104 y/o el dispositivo de descodificación de vídeo 112 pueden estar integrados con un dispositivo de codificación de audio y un dispositivo de descodificación de audio, respectivamente. El dispositivo de codificación de vídeo 104 y/o el dispositivo de descodificación de vídeo 112 también pueden incluir otro hardware o software que sea necesario para implementar las técnicas de codificación descritas anteriormente, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. El dispositivo de codificación de vídeo 104 y el dispositivo de descodificación de vídeo 112 pueden

integrarse como parte de un codificador/descodificador (códec) combinado en un dispositivo respectivo. A continuación, se describe un ejemplo de detalles específicos del dispositivo de codificación 104 con referencia a la FIG. 7. A continuación, se describe un ejemplo de detalles específicos del dispositivo de descodificación 112 con referencia a la FIG. 8.

[0071] Las ampliaciones al estándar HEVC incluyen la ampliación de codificación de vídeo multivista, denominada MV-HEVC, y la ampliación de codificación de vídeo escalable, denominada SHVC. Las ampliaciones MV-HEVC y SHVC comparten el concepto de codificación en capas, con diferentes capas incluidas en el flujo de bits de vídeo codificado. Cada capa en una secuencia de vídeo codificada se direcciona mediante un identificador (ID) de capa único. Un ID de capa puede estar presente en una cabecera de una unidad NAL para identificar una capa con la cual está asociada la unidad NAL. En MV-HEVC, diferentes capas pueden representar diferentes visualizaciones de la misma escena en el flujo de bits de vídeo. En SHVC, se proporcionan diferentes capas escalables que representan el flujo de bits de vídeo en diferentes resoluciones espaciales (o resolución de imagen) o en diferentes fidelidades de reconstrucción. Las capas escalables pueden incluir una capa base (con ID de capa = 0) y una o más capas de mejora (con ID de capa = 1, 2, ..., n). La capa base puede ajustarse a un perfil de la primera versión de HEVC y representa la capa más baja disponible en un flujo de bits. Las capas de mejora tienen una mayor resolución espacial, resolución temporal o velocidad de trama, y/o fidelidad de reconstrucción (o calidad) en comparación con la capa base. Las capas de mejora están organizadas jerárquicamente y pueden (o no) depender de capas inferiores. En algunos ejemplos, las diferentes capas se pueden codificar usando un solo códec estándar (por ejemplo, todas las capas se codifican usando HEVC, SHVC u otra norma de codificación). En algunos ejemplos, diferentes capas pueden codificarse usando un códec de múltiples normas. Por ejemplo, una capa base puede codificarse usando AVC, mientras que una o más capas de mejora pueden codificarse usando las ampliaciones SHVC y/o MV-HEVC de la norma HEVC.

[0072] En general, una capa incluye un conjunto de unidades NAL VCL y un conjunto correspondiente de unidades NAL no VCL. A las unidades NAL se les asigna un valor de ID de capa particular. Las capas pueden ser jerárquicas en el sentido de que una capa puede depender de una capa inferior. Un conjunto de capas se refiere a un conjunto de capas representadas dentro de un flujo de bits que son independientes, lo cual significa que las capas dentro de un conjunto de capas pueden depender de otras capas en el conjunto de capas en el proceso de descodificación, pero no depender de ninguna otra capa para la descodificación. En consecuencia, las capas en un conjunto de capas pueden formar un flujo de bits independiente que puede representar el contenido de vídeo. El conjunto de capas en un conjunto de capas se puede obtener de otro flujo de bits mediante la operación de un proceso de extracción de flujo de bits secundario. Un conjunto de capas puede corresponder al conjunto de capas que se va a descodificar cuando un descodificador desea operar de acuerdo con ciertos parámetros.

[0073] Como se describió previamente, los formatos de medios se pueden usar para encapsular vídeo codificado. Un ejemplo de un formato de medios incluye el formato de archivo de medios de base ISO (ISO/BMFF, especificado en ISO/IEC 14496-12). También existen otros formatos de archivos de medios derivados del ISO/BMFF (ISO/IEC 14496-12), incluido el formato de archivo MPEG-4 (ISO/IEC 14496-14), el formato de archivo 3GPP (3GPP TS 26.244) y el formato de archivo AVC (ISO/IEC 14496-15). Por ejemplo, un flujo de bits de vídeo codificado como se analizó anteriormente puede escribirse o empaquetarse en uno o más archivos usando el ISO/BMFF, un formato de archivo derivado del ISO/BMFF, algún otro formato de archivo y/o una combinación de formatos de archivo que incluye el ISO/BMFF. El archivo o archivos ISO/BMFF se pueden reproducir usando un dispositivo reproductor de vídeo, se pueden transmitir mediante un dispositivo de codificación (o dispositivo de generación de archivos) y luego se pueden mostrar en un dispositivo reproductor, se pueden almacenar y/o se pueden usar en cualquier otro modo adecuado.

[0074] El ISO/BMFF se usa como la base para muchos formatos de encapsulación de códec (por ejemplo, el formato de archivo AVC, entre otros), así como para muchos formatos de contenedor multimedia (por ejemplo, el formato de archivo MPEG-4, el formato de archivo 3GPP (3GP) y el formato de archivo DVB, entre otros). Los medios continuos (por ejemplo, audio y vídeo), los medios estáticos (por ejemplo, imágenes), así como los metadatos se pueden almacenar en un archivo que cumpla con ISO/BMFF. Los archivos estructurados de acuerdo con el ISO/BMFF se pueden usar para muchos propósitos, incluyendo la reproducción local de archivos de medios, la descarga progresiva de un archivo remoto, segmentos para la transmisión en continuo dinámica adaptativa a través de HTTP (DASH), contenedores para contenido que se va a transmitir en continuo e instrucciones de empaquetado para el contenido y el registro de flujos de medios recibidos en tiempo real, entre otros propósitos adecuados. Por ejemplo, aunque originalmente se diseñó para el almacenamiento, el ISO/BMFF ha demostrado ser muy valioso para la transmisión en continuo de medios (por ejemplo, para una descarga progresiva o DASH). Para propósitos de transmisión continua, se pueden usar fragmentos definidos en ISO/BMFF.

[0075] El ISO/BMFF está diseñado para contener la información temporizada en un formato flexible y extensible que facilite el intercambio, la gestión, la edición y la presentación de los medios. La presentación de los medios puede ser "local" para el sistema que contiene la presentación o la presentación puede ser a través de una red u otro mecanismo de entrega de flujo (por ejemplo, DASH u otro mecanismo de flujo en continuo adecuado). Una

"presentación", según se define en la especificación ISOBMFF, es una secuencia de imágenes, a menudo relacionadas por haber sido capturadas secuencialmente por un dispositivo de captura de vídeo, o relacionadas por alguna otra razón. En el presente documento, una presentación también puede denominarse presentación de película o vídeo. Una presentación puede incluir audio. Una sola presentación puede estar contenida en uno o más archivos, conteniendo un archivo los metadatos de toda la presentación. Los metadatos incluyen información como datos de temporización y de tramas, descriptores, punteros, parámetros y otra información que describe la presentación. Los metadatos no incluyen los datos de vídeo y/o audio en sí. Los archivos distintos del archivo que contiene los metadatos no necesitan formatearse de acuerdo con el ISOBMFF, y solo necesitan formatearse de manera que estos archivos puedan ser referenciados por los metadatos.

[0076] La estructura de archivo de un archivo de medios de base ISO está orientada a objetos y la estructura de un objeto individual en el archivo se puede inferir directamente del tipo de objeto. Los objetos en un archivo de medios de base ISO se denominan "cajas" por la especificación ISOBMFF. Un archivo de medios de base ISO está estructurado como una secuencia de cajas, que pueden contener otras cajas. Una caja es la estructura de sintaxis elemental en el ISOBMFF, que incluye un tipo de caja codificada de cuatro caracteres, un recuento de bytes de la caja y la carga útil. Las cajas pueden incluir una cabecera que proporcione un tamaño y un tipo para la caja. El tamaño describe el tamaño completo de la caja, incluida la cabecera, los campos y todas las cajas contenidas dentro de la caja. Las cajas con un tipo que no es reconocido por un dispositivo reproductor típicamente se ignoran y se omiten.

[0077] Un archivo ISOBMFF puede contener diferentes tipos de cajas. Por ejemplo, una caja de película ("moov") contiene los metadatos para los flujos continuos de medios presentes en el archivo, cada uno representado en el archivo como una pista. Los metadatos para una pista están contenidos en una caja de pista ("trak"), mientras que el contenido de medios de una pista está contenido en una caja de datos de medios ("mdat") o directamente en un archivo separado. Puede haber diferentes tipos de pistas. El ISOBMFF especifica, por ejemplo, los siguientes tipos de pistas: una pista de medios, que contiene un flujo de medios elemental, una pista de indicaciones, que incluye instrucciones de transmisión de medios o representa un flujo de paquetes recibido, y una pista de metadatos temporizados, que comprende metadatos sincronizados en el tiempo.

[0078] El contenido de medios para pistas incluya una secuencia de muestras, tales como unas unidades de acceso de audio o vídeo. Los metadatos para cada pista incluyen una lista de entradas de descripción de muestra, cada una de las cuales proporciona el formato de codificación o encapsulación usado en la pista y los datos de inicialización necesarios para procesar ese formato. Cada muestra está asociada a una de las entradas de descripción de muestra de la pista.

[0079] El ISOBMFF permite el uso de metadatos específicos de muestra con diversos mecanismos. Las cajas específicas dentro de la caja de tabla de muestras ("stbl") se han normalizado para responder a necesidades comunes. Por ejemplo, se puede usar una caja de muestras de sincronización ("stss") para enumerar las muestras de acceso aleatorio de la pista. El mecanismo de agrupación de muestras permite la correlación de muestras de acuerdo con un tipo de agrupación de cuatro caracteres en grupos de muestras que comparten la misma propiedad especificada como una entrada de descripción de grupo de muestras en el archivo. Se han especificado diversos tipos de agrupación en el ISOBMFF.

[0080] La FIG. 2 es un diagrama que ilustra un ejemplo de un archivo 200 con una estructura de archivo que sigue el formato de archivo de medios de base ISO. El archivo de medios de base ISO 200 también puede denominarse archivo de formato de medios. Una presentación de medios puede estar contenida, pero no siempre, en un archivo, en cuyo caso la presentación de medios está incluida en el archivo. El archivo 200 incluye un contenedor 202 de película (o "caja de película"). El contenedor 202 de película puede contener los metadatos de los medios, que pueden incluir, por ejemplo, una o más pistas de vídeo y una o más pistas de audio. Por ejemplo, una pista 210 de vídeo puede contener información sobre varias capas de un vídeo, que pueden almacenarse en uno o más contenedores 214 de información de medios. Por ejemplo, un contenedor 214 de información de medios puede incluir una tabla de muestras, que proporciona información sobre las muestras de vídeo para el vídeo. En varias implementaciones, el fragmento 222 de datos de vídeo y el fragmento 224 de datos de audio están contenidos en el contenedor 204 de datos de medios. En algunas implementaciones, el fragmento 222 de datos de vídeo y los datos de audio 224 pueden estar contenidos en uno o más archivos (distintos del archivo 200).

[0081] En varias implementaciones, una presentación (por ejemplo, una secuencia de movimiento) puede estar contenida en varios archivos. Toda la información de temporización y de tramas (posición y tamaño) puede estar en el archivo de medios de base ISO y los archivos auxiliares pueden usar fundamentalmente cualquier formato.

[0082] Los archivos ISO tienen una estructura lógica, una estructura de tiempo y una estructura física. No es necesario acoplar las diferentes estructuras. La estructura lógica del archivo es de una película que a su vez contiene un conjunto de pistas paralelas en el tiempo (por ejemplo, la pista 210 de vídeo). La estructura temporal del archivo es que las pistas contengan secuencias de muestras en el tiempo, y esas secuencias se correlacionen con la línea de tiempo de toda la película mediante listas de edición opcionales.

[0083] La estructura física del archivo separa los datos necesarios para la descomposición lógica, temporal y estructural de las propias muestras de datos de medios. Esta información estructural puede concentrarse en una caja de película (por ejemplo, un contenedor 202 de película), posiblemente extendida en el tiempo por cajas de fragmentos de película. La caja de película documenta las relaciones lógicas y de temporización de las muestras, y también contiene punteros a donde están ubicadas las muestras. Los punteros pueden apuntar al mismo archivo o a otro archivo, al que se puede hacer referencia, por ejemplo, mediante un localizador uniforme de recursos (URL) o similar.

[0084] Cada flujo de medios está contenido en una pista especializada para ese tipo de medio. Por ejemplo, en el ejemplo ilustrado en la FIG. 2, el contenedor 202 de película incluye una pista 210 de vídeo y una pista 216 de audio. El contenedor 202 de película también puede incluir una pista 218 de indicaciones, que puede incluir instrucciones de transmisión desde la pista 210 de vídeo y/o la pista 216 de audio, o puede representar otra información sobre otras pistas en el contenedor 202 de película u otro contenedor de película (no mostrado) del archivo 200. Cada pista puede parametrizarse adicionalmente mediante una entrada de muestra. Por ejemplo, en el ejemplo ilustrado, la pista 210 de vídeo incluye un contenedor 214 de información de medios, que incluye una tabla de muestras (denominada "tabla de muestras"). Una entrada de muestras contiene el "nombre" del tipo exacto de medios (por ejemplo, el tipo de descodificador necesario para descodificar el flujo) y cualquier parametrización necesaria para ese descodificador. El nombre puede tomar la forma de un código de cuatro caracteres (por ejemplo, moov, trak u otro código de nombre adecuado). Hay formatos de entrada de muestras definidos para varios tipos de medios. Las entradas de muestras pueden incluir, además, punteros a fragmentos de datos de vídeo (por ejemplo, los fragmentos 222 de datos de vídeo) en una caja 220 en el contenedor 204 de datos de medios. La caja 220 incluye muestras de vídeo entrelazadas, ordenadas en el tiempo (organizadas en fragmentos de datos de vídeo, tales como los fragmentos de datos de vídeo 222), tramas de audio (por ejemplo, en el fragmento 224 de datos de audio) e instrucciones de indicaciones (por ejemplo, en el fragmento 226 de instrucciones de indicaciones).

[0085] El soporte para metadatos puede adoptar diferentes formas. En un ejemplo, los metadatos temporizados pueden almacenarse en una pista apropiada y pueden sincronizarse según se desee con los datos de medios descritos por los metadatos. En un segundo ejemplo, hay soporte general para metadatos no temporizados adjuntos a la película o a una pista individual. El soporte estructural es general y permite, como en los datos de medios, el almacenamiento de recursos de metadatos en otra parte del archivo o en otro archivo.

[0086] En algunos casos, una pista de un archivo de vídeo puede contener varias capas. La pista de vídeo también puede incluir una cabecera de pista (por ejemplo, la cabecera 212 de pista), que puede contener alguna información sobre el contenido de la pista de vídeo (por ejemplo, la pista 210 de vídeo). Por ejemplo, la cabecera de la pista puede incluir una caja de información de contenido de la pista (también denominada "tcon"). La caja tcon puede enumerar todas las capas y subcapas en la pista de vídeo. Un archivo de vídeo también puede incluir una caja de información del punto de operación (también denominada caja 'oinf'). La caja oinf registra información sobre los puntos de operación, como las capas y subcapas que constituyen el punto de operación, las dependencias (si las hay) entre los puntos de operación, el perfil, nivel y parámetro de nivel del punto de operación, y otra información pertinente de dichos puntos de operación. En algunos casos, un punto de operación también puede denominarse punto de operación.

[0087] La FIG. 3 es un diagrama que ilustra otro ejemplo de un archivo de medios de base ISO 300 formateado de acuerdo con el ISOBMFF. El archivo de medios de base ISO 300 también puede denominarse archivo de formato de medios. El archivo de medios de base ISO 300 contiene datos y metadatos para una presentación de vídeo. En el nivel superior del archivo 300, se incluye una caja 310 de tipo de archivo, una caja 320 de película y uno o más fragmentos 330a, 330b, 330c, a 330n. Otras cajas que se pueden incluir en este nivel pero que no están representadas en este ejemplo incluyen cajas de espacio libre, cajas de metadatos y cajas de datos de medios, entre otras.

[0088] La caja 310 de tipo de archivo se identifica mediante el tipo de caja "ftyp". La caja 310 de tipo de archivo típicamente se coloca lo antes posible en el archivo de medios de base ISO 300. La caja 310 de tipo de archivo identifica una especificación ISOBMFF que es la más adecuada para analizar el archivo. "La mayoría" en este caso significa que el archivo de medios de base ISO 300 puede haber sido formateado de acuerdo con una especificación ISOBMFF particular, pero es probable que sea compatible con otras iteraciones de la especificación. Esta especificación más adecuada se conoce como la marca principal. Un dispositivo reproductor puede usar la marca principal para determinar si el dispositivo es capaz de descodificar y mostrar el contenido del archivo. La caja 310 de tipo de archivo también puede incluir un número de versión, que se puede usar para indicar una versión de la especificación ISOBMFF. La caja 310 de tipo de archivo también puede incluir una lista de marcas compatibles, que incluye una lista de otras marcas con las que el archivo es compatible. Un archivo de medios de base ISO puede ser compatible con más de una marca importante.

[0089] Cuando un archivo de medios de base ISO incluye una caja de tipo de archivo (como el archivo de medios de base ISO 300), solo hay una caja de tipo de archivo. En algunos casos, un archivo de medios de base

ISO puede omitir la caja de tipo de archivo para que sea compatible con dispositivos de reproducción más antiguos. Cuando un archivo de medios de base ISO no incluye una caja de tipo de archivo, un dispositivo reproductor puede asumir una marca principal predeterminada (por ejemplo, mp41), una versión secundaria (por ejemplo, "0") y una marca compatible (por ejemplo, mp41, isom, iso2, avc1 o similares).

[0090] El archivo de medios de base ISO 300 incluye, además, una caja 320 de película, que contiene los metadatos para la presentación. La caja 320 de película se identifica por el tipo de caja "moov". ISO/IEC 14496-12 establece que una presentación, ya sea contenida en un archivo o en varios archivos, puede incluir solo una caja 320 de película. Con frecuencia, la caja de película está cerca del comienzo de un archivo de medios de base ISO (por ejemplo, como se indica por la ubicación de la caja 320 de película en el archivo de medios base ISO 300). La caja 320 de película incluye una caja 322 de cabecera de película y puede incluir una o más cajas 324 de pistas, así como otras cajas.

[0091] La caja 322 de cabecera de película, identificada por el tipo de caja "mvhd", puede incluir información que es independiente de los medios y pertinente para la presentación en su conjunto. Por ejemplo, la caja 322 de cabecera de película puede incluir información tal como un tiempo de creación, un tiempo de modificación, una escala de tiempo y/o una duración de la presentación, entre otras cosas. La caja 322 de cabecera de película también puede incluir un identificador que identifica la siguiente pista en la presentación. Por ejemplo, el identificador puede apuntar a la caja 324 de pistas contenida por la caja 320 de película en el ejemplo ilustrado.

[0092] La caja 324 de pistas, identificada por el tipo de caja "trak", puede contener la información de una pista para una presentación. Una presentación puede incluir una o más pistas, donde cada pista es independiente de otras pistas en la presentación. Cada pista puede incluir la información temporal y espacial que es específica del contenido de la pista, y cada pista se puede asociar con una caja de medios. Los datos en una pista pueden ser datos de medios, en cuyo caso la pista es una pista de medios, o los datos pueden ser información de empaquetado para protocolos de transmisión en continuo, en cuyo caso la pista es una pista de indicaciones. Los datos de medios incluyen, por ejemplo, datos de vídeo y audio. En el ejemplo mostrado en la FIG. 3, la caja 324 de pistas de ejemplo incluye una caja 324a de cabecera de pista y una caja 324b de medios. Una caja de pistas puede incluir otras cajas, como una caja de referencia de pistas, una caja de grupo de pistas, una caja de edición, una caja de datos de usuario, una caja de meta, entre otras.

[0093] La caja 324a de cabecera de pista, identificada por el tipo de caja "tkhd", puede especificar las características de una pista contenida en la caja 324 de pistas. Por ejemplo, la caja 324a de cabecera de pista puede incluir un tiempo de creación, un tiempo de modificación, duración, identificador de pista, identificador de capa, identificador de grupo, volumen, ancho y/o altura de la pista, entre otras cosas. Para una pista de medios, la caja 324a de cabecera de pista puede identificar, además, si la pista está habilitada, si la pista debiera reproducirse como parte de la presentación, si la pista puede usarse para previsualizar la presentación, entre otros usos de la pista. En general, se supone que la presentación de una pista está al comienzo de una presentación. La caja 324 de pistas puede incluir una caja de lista de edición (no mostrada) que puede incluir un mapa de línea de tiempo explícito. El mapa de la línea de tiempo puede especificar, entre otras cosas, un tiempo de desfase para la pista, donde el desfase indica una hora de inicio, después del comienzo de la presentación, para la pista.

[0094] En el ejemplo ilustrado, la caja 324 de pistas también incluye una caja 324b de medios, identificada por el tipo de caja "mdia". La caja 324b de medios puede contener los objetos y la información sobre los datos de medios en la pista. Por ejemplo, la caja 324b de medios puede contener una caja de referencia del manejador, que puede identificar el tipo de medio de la pista y el proceso mediante el cual se presentan los medios en la pista. Como otro ejemplo, la caja 324b de medios puede contener una caja de información de medios, que puede especificar las características de los medios en la pista. La caja de información de medios puede incluir, además, una tabla de muestras, como se describe anteriormente con respecto a la FIG. 2, donde cada muestra describe un fragmento de datos de medios (por ejemplo, datos de vídeo o audio) que incluyen, por ejemplo, la ubicación de los datos para la muestra. Los datos de una muestra se almacenan en una caja de datos de medios, que se analiza más adelante. Como ocurre con la mayoría de las otras cajas, la caja 324b de medios también puede incluir una caja de cabecera de medios.

[0095] En el ejemplo ilustrado, el archivo de medios de base ISO 300 de ejemplo también incluye múltiples fragmentos 330a, 330b, 330c, a 330n de la presentación. Los fragmentos 330a, 330b, 330c, a 330n no son cajas ISOBMFF, sino que más bien describen una caja 332 de fragmentos de película y una caja 338 de datos de medios a la que hace referencia la caja 332 de fragmentos de película. La caja 332 de fragmentos de película y la caja 338 de datos de medios son cajas de nivel superior, pero están agrupadas aquí para indicar la relación entre una caja 332 de fragmentos de película y una caja 338 de datos de medios.

[0096] Una caja 332 de fragmentos de película, identificada por el tipo de caja "moof", puede ampliar una presentación al incluir información adicional que de otro modo se almacenaría en la caja 320 de películas. Usando cajas de fragmentos de película (por ejemplo, la caja 332 de fragmentos de película), se puede crear una presentación de forma incremental. Una caja 332 de fragmentos de película puede incluir una caja 334 de

cabecera de fragmentos de película y una caja 336 de fragmentos de pista, así como otras cajas no ilustradas aquí.

[0097] La caja 334 de cabecera de fragmentos de película, identificada por el tipo de caja "mfhd", puede incluir un número de secuencia. Un dispositivo reproductor puede usar el número de secuencia para verificar que el fragmento 330a incluye el siguiente dato para la presentación. En algunos casos, el contenido de un archivo, o los archivos de una presentación, se pueden proporcionar a un dispositivo reproductor sin ordenar. Por ejemplo, los paquetes de red pueden llegar con frecuencia en un orden distinto al orden en que se transmitieron originalmente. En estos casos, el número de secuencia puede ayudar al dispositivo reproductor a determinar el orden correcto de los fragmentos.

[0098] La caja 332 de fragmentos de película también puede incluir una o más cajas 336 de fragmentos de pista, identificadas por el tipo de caja "traf". Una caja 332 de fragmentos de película puede incluir un conjunto de fragmentos de pista, cero o más por pista. Los fragmentos de pista pueden contener cero o más ejecuciones de pista, de las que cada una describe una ejecución contigua de muestras para esa pista. Los fragmentos de pista se pueden usar para agregar tiempo vacío a una pista, además de agregar muestras a la pista.

[0099] La caja 338 de datos de medios, identificada por el tipo de caja "mdat", contiene datos de medios. En las pistas de vídeo, la caja 338 de datos de medios puede contener tramas de vídeo, unidades de acceso, unidades NAL u otra forma de datos de vídeo. Una caja de datos de medios puede incluir de forma alternativa o adicional datos de audio. Una presentación puede incluir cero o más cajas de datos de medios, contenidos en uno o más archivos individuales. Los datos de medios se describen mediante metadatos. En el ejemplo ilustrado, los datos de medios en la caja 338 de datos de medios pueden describirse mediante metadatos incluidos en la caja 336 de fragmentos de pista. En otros ejemplos, los datos de medios en una caja de datos de medios pueden describirse mediante metadatos en la caja 320 de película. Los metadatos pueden referirse a datos de medios particulares mediante un desfase absoluto dentro del archivo 300, de modo que se puede omitir una cabecera de datos de medios y/o espacio libre dentro de la caja 338 de datos de medios.

[0100] Otros fragmentos 330b, 330c, a 330n en el archivo de medios de base ISO 300 pueden contener cajas similares a las ilustradas para el primer fragmento 330a, y/o pueden contener otras cajas.

[0101] Como se señaló anteriormente, el ISOBMFF incluye soporte para la transmisión en continuo de datos de medios a través de una red, además de admitir la reproducción local de los medios. El archivo o archivos que incluyen una presentación de película pueden incluir pistas adicionales, llamadas pistas de indicaciones, que contienen instrucciones que pueden ayudar a un servidor de transmisión en continuo a formar y transmitir el archivo o archivos como paquetes. Estas instrucciones pueden incluir, por ejemplo, datos para que el servidor envíe (por ejemplo, información de cabecera) o referencias a segmentos de los datos de los medios. Un archivo puede incluir pistas de indicaciones separadas para diferentes protocolos de transmisión en continuo. Las pistas de indicaciones también se pueden agregar a un archivo sin necesidad de reformatear el archivo.

[0102] Un procedimiento para la transmisión en continuo de datos de medios es la transmisión dinámica adaptativa sobre el protocolo de transferencia de hipertexto (HTTP) o DASH (definido en ISO/IEC 23009-1:2014). DASH, que también se conoce como MPEG-DASH, es una técnica de transmisión de velocidad de transferencia de bits adaptativa que permite la transmisión en continuo de contenido de medios de alta calidad mediante servidores web HTTP convencionales. DASH opera dividiendo el contenido de los medios en una secuencia de pequeños segmentos de archivos basados en HTTP, donde cada segmento contiene un intervalo de tiempo corto del contenido. Con DASH, un servidor puede proporcionar el contenido de medios a diferentes velocidades de transferencia de bits. Un dispositivo cliente que está reproduciendo los medios puede seleccionar entre las velocidades de transferencia de bits alternativas al descargar un siguiente segmento y, por lo tanto, adaptarse a las condiciones cambiantes de la red. DASH utiliza la infraestructura del servidor web HTTP de Internet para entregar contenido a través de la World Wide Web. DASH es independiente del códec usado para codificar y descodificar el contenido de medios y, por lo tanto, opera con códecs como H.264 y HEVC, entre otros.

[0103] El ISOBMFF incluye un diseño de esquema restringido que se usa cuando un creador (por ejemplo, un autor, productor o similar) del archivo de formato de medios requiere que el reproductor o el renderizador realice ciertas acciones al procesar contenido de medios asociado con el archivo de medios de base ISO. El esquema restringido permite a los reproductores inspeccionar el archivo de formato de medios para determinar fácilmente dichos requisitos para renderizar el contenido de medios y puede evitar que los reproductores heredados descarguen, descodifiquen y rendericen archivos que requieran procesamiento adicional. Un reproductor heredado puede incluir un dispositivo de reproducción que no está diseñado para procesar cierto contenido de medios porque, por ejemplo, el reproductor solo puede reproducir contenido de vídeo codificado sin el procesamiento requerido. El esquema restringido puede aplicarse cuando el contenido solo debe ser descodificado por reproductores que lo presentan correctamente. Ejemplos ilustrativos de tipos de esquemas restringidos incluyen un tipo de esquema restringido para disposiciones de empaquetado de tramas de vídeo estereoscópico (denotado como 'stvi'), un tipo de esquema restringido para vídeo de realidad virtual (denotado como 'vrvid'), un tipo de esquema restringido para vídeo omnidireccional proyectado (denotado como 'podv'), un

tipo de esquema restringido para vídeo proyectado equirrectangular (denotado como 'erpv'), un tipo de esquema restringido para vídeo proyectado equirrectangular o cubemap empaquetado (denotado como 'ercm'), un tipo de esquema restringido para vídeo omnidireccional de ojo de pez (denotado como 'fodv'), entre otros.

5 **[0104]** Los aspectos del mecanismo del esquema restringido son similares a la transformación de protección de contenido, donde las entradas de muestras se ocultan detrás de entradas de muestras genéricas, 'encv', 'enca' y similares, lo que indica medios cifrados o encapsulados. El mecanismo análogo para el vídeo restringido usa una transformación con la entrada de muestras genérica 'resv'.

10 **[0105]** Un ejemplo de aplicación del esquema restringido se aplica de la siguiente manera:

1) El código de cuatro caracteres de la entrada de muestras se reemplaza por un nuevo código de entrada de muestras 'resv' (que significa vídeo restringido).

15 2) Se agrega una caja de información de esquema restringido a la descripción de la muestra, dejando todas las demás cajas sin modificar.

3) El tipo de entrada de muestras original se almacena dentro de una caja de formato original contenida en la caja de información de esquema restringido.

20 **[0106]** Una caja de información de esquema restringido (denotada como RestrictedSchemeInfoBox) tiene el mismo formato que una caja de información de esquema de protección (denotada como ProtectionSchemeInfoBox, excepto por que la caja de información de esquema restringido usa el identificador 'rinf' en lugar de 'sinf' (como se muestra en los ejemplos a continuación)). El tipo de entrada de muestras original se encuentra en la caja de formato original ubicada en la caja de información de esquema restringido (de la misma manera que la caja de información de esquema de protección es para los medios cifrados).

25 **[0107]** La naturaleza exacta de la restricción se define en una caja superior de esquema (denominada SchemeTypeBox) contenida en la caja de información de esquema restringido, y los datos necesarios para ese esquema se almacenan en una caja de información de esquema (denominada SchemeInformationBox).

30 **[0108]** Los esquemas de restricción y protección se pueden aplicar al mismo tiempo. El orden de las transformaciones se deriva del código de cuatro caracteres de la entrada de muestras. Por ejemplo, si el tipo de entrada de muestras es 'resv', deshacer la transformación del esquema restringido puede dar como resultado un tipo de entrada de muestras 'encv', lo que indica que los medios están protegidos. Se puede deshacer una transformación (por ejemplo, un esquema de protección, un esquema de protección o similar) mediante un proceso de análisis analizando la estructura de archivo del archivo de medios de base ISO para desenvolver el archivo capa por capa para obtener y procesar la información en el archivo. Por ejemplo, después de analizar la transformación de esquema restringido para el tipo de entrada de muestras de vídeo restringido 'resv' para obtener información relacionada con un tipo de esquema particular usado para el tipo de entrada de muestras 'resv' (por ejemplo, un tipo de esquema restringido para disposiciones de vídeo estereoscópico, un tipo de esquema restringido para vídeo de realidad virtual, u otro tipo de esquema), el dispositivo reproductor encontrará información en el archivo relacionado con el tipo de entrada de muestras 'encv'. El dispositivo reproductor puede entonces analizar el contenido relacionado con el cifrado para obtener la información necesaria para descifrar el contenido. Aún más, después de analizar y procesar la información de cifrado, el dispositivo reproductor puede encontrar la información de codificación necesaria para decodificar el contenido. A continuación, se puede procesar el diferente contenido leído del archivo. Por ejemplo, el contenido de vídeo se puede descifrar usando la información relacionada con el cifrado del archivo. El vídeo descifrado se puede decodificar usando la información relacionada con la codificación en el archivo, y el vídeo decodificado se puede reproducir usando la información de esquema restringido (por ejemplo, desempaquetar una trama usando información de tipo de trama 'stvi' para disposiciones de vídeo estereoscópico) junto con otra información en el archivo.

45 **[0109]** Tenga en cuenta que si el autor del archivo solo desea proporcionar información de advertencia sin impedir que los reproductores heredados reproduzcan el archivo, la caja de información de esquema restringido se puede colocar dentro de la entrada de muestras sin aplicar la transformación de código de cuatro caracteres. En este caso, no es necesario incluir una caja de formato original.

50 **[0110]** La caja de información de esquema restringido se define actualmente de la siguiente manera:

60 Caja de información de esquema restringido

Tipos de cajas:	'rinf'
Contenedor:	Entrada de muestras restringidas o entrada de muestras
Obligatorio:	Sí

Cantidad: Exactamente una

La caja de información de esquema restringido contiene toda la información necesaria para comprender el esquema restringido aplicado y sus parámetros. También documenta el tipo de entrada de muestras original (sin transformar) de los medios. La caja de información de esquema restringido es una caja contenedor. Es obligatorio en una entrada de muestras que usa un código que indica un flujo restringido, es decir, 'resv'. Cuando se usa en una entrada de muestras restringidas, esta caja debe contener la caja de formato original para documentar el tipo de entrada de muestras original y una caja de tipo de esquema. Es posible que se requiera una caja de información de esquema según el esquema de restricción.

```
aligned(8) class RestrictedSchemeInfoBox (fmt) extends Box ('rinf') {
    OriginalFormatBox(fmt)    original_format;
    SchemeTypeBox             scheme_type_box;
    SchemeInformationBox      info;                //
opcional
}
```

[0111] La caja de tipo de esquema se define actualmente de la siguiente manera:

Caja de tipo de esquema

Tipos de cajas: 'schm'

Contenedor: Caja de información de esquema de protección ('sinf'), caja de información de esquema restringido ('rinf') o caja de proceso SRTP ('srpp')

Obligatorio: No

Cantidad: Cero o una en 'sinf', dependiendo de la estructura de protección; Exactamente una en 'rinf' y 'srpp'

[0112] La caja de tipo de esquema ('schm') identifica el esquema de protección o restricción.

```
aligned(8) class SchemeTypeBox extends FullBox('schm', 0, flags) {
    unsigned int(32) scheme_type;    // 4CC identificando el
esquema
    unsigned int(32) scheme_version; // versión del esquema
    if (flags & 0x000001) {
        unsigned int(8)    // uri del navegador
        scheme_uri[];
    }
}
```

scheme_type es el código que define el esquema de protección o restricción. scheme_version es la versión del esquema (usado para crear el contenido). scheme_URI permite la opción de dirigir al usuario a una página web si no tiene el esquema instalado en su sistema. Es un URI absoluto formado como una cadena terminada en nulo en caracteres UTF-8.

[0113] La caja de información de esquema se define actualmente de la siguiente manera:

Caja de información de esquema

Tipos de cajas: 'schi'

Contenedor: Caja de información de esquema de protección ('sinf'), caja de información de esquema restringido ('rinf') o caja de proceso SRTP ('srpp')

Obligatorio: No

Cantidad: Cero o una

[0114] La caja de información de esquema es una caja contenedor que solo es interpretada por el esquema que se está usando. Toda la información que necesite el sistema de cifrado o restricción se almacena aquí. El

contenido de esta caja es una serie de cajas cuyo tipo y formato están definidos por el esquema declarado en la caja de tipo de esquema.

```
aligned(8) class SchemeInformationBox extends Box('schi') {
    Box scheme_specific_data[];
}
```

[0115] Existen varios problemas con respecto al esquema restringido en el ISOBMFF. Por ejemplo, en el diseño existente del esquema restringido, solo puede haber una caja de información de esquema restringido en su contenedor (el contenedor de entrada de muestras restringidas o el contenedor de entrada de muestras). Además, solo puede haber una caja de tipo de esquema en la caja de información de esquema restringido, y solo puede haber cero o una caja de información de esquema en la caja de información de esquema restringido.

[0116] Tales restricciones no permiten que se apliquen simultáneamente varios tipos de esquemas restringidos al mismo contenido de medios asociado con una entrada de muestras. Una entrada de muestras se puede asociar con una o más muestras de un archivo de vídeo. Sin embargo, existen muchas aplicaciones de medios del mundo real que utilizan contenido de vídeo al que se aplican múltiples esquemas restringidos. Usando vídeo de realidad virtual estereoscópico empaquetado en tramas como un ejemplo ilustrativo, las restricciones en el esquema restringido evitan la posibilidad de aplicar tanto el esquema restringido para disposiciones de vídeo estereoscópico (denotado como 'stvi' y definido en la cláusula 8.15.4 del ISOBMFF) como el esquema restringido para vídeo de realidad virtual (denotado como 'vrvd' en el presente documento y definido en el documento de entrada MPEG m37914) al mismo contenido de vídeo asociado con una entrada de muestras. Existen muchos otros ejemplos en los que se pueden aplicar múltiples esquemas restringidos al mismo contenido de medios asociado con una entrada de muestras.

[0117] Los sistemas y técnicas se describen en el presente documento para abordar los problemas señalados anteriormente con el esquema restringido. Algunos de los aspectos descritos en el presente documento se pueden aplicar de forma independiente y algunos de ellos se pueden aplicar en combinación. En algunos casos, se puede generar un archivo de formato de medios (de acuerdo con un formato de archivo de medios, como ISOBMFF u otro formato adecuado) que incluye múltiples esquemas restringidos que son aplicables a cierto contenido de vídeo. Por ejemplo, los múltiples esquemas restringidos se pueden aplicar simultáneamente al mismo contenido de medios asociado con una entrada de muestras. El diseño mejorado se aplica a cualquier tipo de códec de vídeo. Un dispositivo reproductor puede recibir un archivo de medios con múltiples esquemas restringidos indicados en el mismo, y puede analizar el archivo para obtener la información necesaria para aplicar los múltiples esquemas restringidos al contenido de vídeo. Por ejemplo, usando la información del archivo, el dispositivo reproductor puede aplicar los múltiples esquemas restringidos al contenido de medios asociado con una entrada de muestras en el archivo.

[0118] En algunos ejemplos, se pueden incluir cajas de tipos de esquema múltiples en una caja de información de esquema restringido contenida en una entrada de muestras de un archivo de formato de medios. Como se señaló anteriormente, una caja de tipo de esquema define la naturaleza exacta de una restricción particular. Por ejemplo, una caja de tipo de esquema identifica un tipo de esquema particular que es aplicable al contenido de vídeo (por ejemplo, el tipo de esquema al que se ajusta la pista) asociado con una entrada de muestras. Las cajas de tipos de esquemas múltiples proporcionan la capacidad de identificar una pluralidad de tipos de esquemas restringidos requeridos para procesar el contenido de medios asociado con la entrada de muestras. En algunos casos, una caja de tipo de esquema también puede denominarse caja de tipo de esquema compatible. En algunos ejemplos, también se pueden incluir cajas de información de esquemas múltiples en la caja de información de esquema restringido además de las cajas de tipos de esquema múltiples. Los datos necesarios para un tipo de esquema particular se almacenan en una caja de información de esquema correspondiente. Por ejemplo, se puede incluir una caja de información de esquema en la caja de información de esquema restringido para cada caja de tipo de esquema. Una caja de información de esquema restringido (restrictedSchemeInfoBox) se basa en Box() en lugar de FullBox(), no hay un campo de versión, por lo que no se puede definir una nueva versión de la caja. Sin embargo, este procedimiento es compatible con versiones anteriores, ya que el diseño permite agregar más cajas opcionales al final de una caja.

[0119] A continuación, se proporciona un ejemplo ilustrativo de un modo de realización detallado de un archivo de formato de medios que especifica múltiples esquemas restringidos (usando cajas de tipos de esquema múltiples y cajas de información de esquema) que son aplicables al contenido de medios asociado con una entrada de muestras. En el siguiente ejemplo, el texto adicional agregado a las implementaciones descritas en el diseño de esquema restringido de las normas ISO/IEC (por ejemplo, ISO/IEC 14496-n, ISO/IEC 15444-n o similares) se muestra entre corchetes <*> (por ejemplo, como <"texto adicional">).

[0120] En el ejemplo ilustrativo, para implementar un archivo de formato de medios que especifica múltiples esquemas restringidos aplicables al contenido de medios asociado con una entrada de muestras, las definiciones de la caja de información de esquema restringido, la caja de tipo de esquema y la caja de información de esquema se cambian como se muestra a continuación:

Caja de información de esquema restringido

Tipos de cajas:	'rinf'
Contenedor:	Entrada de muestras restringidas o entrada de muestras
Obligatorio:	Sí
Cantidad:	Exactamente una

5 **[0121]** La caja de información de esquema restringido contiene toda la información necesaria para comprender el esquema de restricción aplicado y sus parámetros. También documenta el tipo de entrada de muestras original (sin transformar) de los medios. La caja de información de esquema restringido es una caja contenedor. Es obligatorio en una entrada de muestras que usa un código que indica un flujo restringido, es decir, 'resv'.

10 **[0122]** Cuando se usa en una entrada de muestras restringidas, esta caja debe contener la caja de formato original para documentar el tipo de entrada de muestras original y una caja de tipo de esquema. Es posible que se requiera una caja de información de esquema según el esquema de restricción.

```
aligned(8) class RestrictedSchemeInfoBox(fmt) extends Box ('rinf') {
    OriginalFormatBox (fmt) original format;
    SchemeTypeBox<#1>                                scheme type box<#1>;
    SchemeInformationBox<#1> optional <#1>;           //
opcional
    <...
    SchemeTypeBox#i                                scheme type box#i;           //
opcional
    SchemeInformationBox#i      opcional#i;
opcional
    ...
    SchemeTypeBox#N                                scheme_type_box#N;           //
opcional
    SchemeInformationBox#N optional#N;
    //opcional>
}
```

15

Caja de tipo de esquema

Tipos de 'schm'
cajas:

Contenedor: Caja de información de esquema de protección ('sinf'), caja de información de esquema restringido ('rinf') o caja de proceso SRTP ('srpp')

Obligatorio: No

Cantidad: Cero o una en 'sinf', dependiendo de la estructura de protección; Exactamente una en 'srpp'; <Una o más en 'rinf'>

20 **[0123]** La caja de tipo de esquema ('schm') identifica el tipo de esquema de protección o restricción.

```
aligned(8) class SchemeTypeBox extends FullBox('schm', 0, flags) {
    unsigned int(32) scheme_type;           // 4CC identificando el
esquema
    unsigned int(32) scheme_version;         // versión de esquema
    if (flags & 0x000001) {
        unsigned int (8) scheme_uri[];      // uri del navegador
    }
}
```

25 scheme_type es el código que define el esquema de protección o restricción. scheme_version es la versión del esquema (usado para crear el contenido) scheme_URI permite la opción de dirigir al usuario a una página web si no tiene el esquema instalado en su sistema. Es un URI absoluto formado como una cadena terminada en nulo en caracteres UTF-8.

Caja de información de esquema

Tipos de
cajas: 'schi'

Contenedor: Caja de información de esquema de protección ('sinf'), caja de información de esquema restringido ('rinf') o caja de proceso SRTP ('srpp')

Obligatorio: No

Cantidad: Cero o una en 'sinf' y 'srpp'; <Cero o más en 'rinf'>

[0124] La caja de información de esquema es una caja contenedor que solo es interpretada por el esquema que se está usando. Toda la información que necesite el sistema de cifrado o restricción se almacena aquí. El contenido de esta caja es una serie de cajas cuyo tipo y formato están definidos por el esquema declarado en la caja de tipo de esquema.

```
aligned(8) class SchemeInformationBox extends Box('schi') {
    Box scheme_specific_data[];
}
```

[0125] Como se muestra en el ejemplo ilustrativo anterior, la caja de información de esquema restringido incluye cajas de tipos de esquema múltiples (SchemeTypeBox n.º 1, SchemeTypeBox n.º i, hasta SchemeTypeBox n.º N), que indican múltiples tipos de esquema que son necesarios para procesar el contenido de medios asociado con el contenedor de entrada de muestras particular que contiene la caja de información de esquema restringido. En algunos casos, cada una de las cajas de tipo de esquema también se puede denotar como CompatibleSchemeTypeBox. Además, la caja de información de esquema restringido incluye múltiples cajas de información de esquemas (SchemeInformationBox n.º 1, SchemeInformationBox n.º i, hasta SchemeInformationBox n.º N), que incluyen los datos necesarios para realizar el tipo de esquema correspondiente. Por ejemplo, SchemeInformationBox n.º 1 puede incluir una o más cajas que incluyen los datos necesarios para implementar el tipo de esquema identificado por SchemeTypeBox n.º 1.

[0126] En algunos ejemplos, un archivo de formato de medios generado de acuerdo con las técnicas descritas en el presente documento puede incluir múltiples cajas de información de esquema restringido en un contenedor de entrada de muestras para indicar múltiples esquemas restringidos que son aplicables al contenido de medios asociado con esa entrada de muestras. A continuación, se proporciona un ejemplo ilustrativo de un modo de realización detallado de tal técnica. En el siguiente ejemplo, el texto adicional agregado a las implementaciones descritas en el diseño de esquema restringido de las normas ISO/IEC (por ejemplo, ISO/IEC 14496-n, ISO/IEC 15444-n o similares) se muestra entre corchetes <*> (por ejemplo, como <"texto adicional">).

[0127] Las definiciones de la caja de información de esquema restringido se modifican como se muestra a continuación:

Caja de información de esquema restringido

Tipos de cajas: 'rinf'

Contenedor: Entrada de muestras restringidas o entrada de muestras

Obligatorio: Sí

Cantidad: <Una o más>

[0128] La caja de información de esquema restringido contiene toda la información necesaria para comprender el esquema de restricción aplicado y sus parámetros. También documenta el tipo de entrada de muestras original (sin transformar) de los medios. La caja de información de esquema restringido es una caja contenedor. Es obligatorio en una entrada de muestras que usa un código que indica un flujo restringido, es decir, 'resv'.

[0129] Cuando se usa en una entrada de muestras restringidas, esta caja debe contener la caja de formato original para documentar el tipo de entrada de muestras original y una caja de tipo de esquema. Es posible que se requiera una caja de información de esquema según el esquema de restricción.

```
aligned(8) class RestrictedSchemeInfoBox (fmt) extends Box ('rinf') {
    OriginalFormatBox(fmt)          original_format;
    SchemeTypeBox                   scheme_type_box;
    SchemeInformationBox             info;
opcional
}
```

[0130] Como se muestra en el ejemplo ilustrativo anterior, generar un archivo de formato de medios con múltiples cajas de información de esquema restringido en un contenedor de entrada de muestras del archivo de

formato de medios permite que el archivo indique múltiples esquemas restringidos que son aplicables al contenido de medios asociado con la entrada de muestras.

[0131] La FIG. 4 ilustra un ejemplo de una caja 424 de medios que se puede incluir en un archivo de medios de base ISO que especifica múltiples esquemas restringidos (usando múltiples cajas de tipo de esquema 462a, 462b y múltiples cajas de información de esquema 464a, 464b) que son aplicables al contenido de medios asociado una entrada de muestras. Como se analizó anteriormente, una caja de medios se puede incluir en una caja de pistas y puede contener objetos e información que describen datos de medios en la pista. En el ejemplo ilustrado, la caja 424 de medios incluye una caja 442 de información de medios. La caja 424 de medios también puede incluir otras cajas, que no se ilustran aquí.

[0132] La caja 442 de información de medios puede contener objetos que describen información característica de los medios de la pista. Por ejemplo, la caja 442 de información de medios puede incluir una caja de información de datos, que describe la ubicación de la información de medios en la pista. Como otro ejemplo, la caja 442 de información de medios puede incluir una cabecera de medios de vídeo, cuando la pista incluye datos de vídeo. La cabecera de medios de vídeo puede contener información de presentación general que es independiente de la codificación de los medios de vídeo. La caja 442 de información de medios también puede incluir una cabecera de medios de sonido cuando la pista incluye datos de audio.

[0133] La caja 442 de información de medios incluye también una caja 444 de tabla de muestras. La caja 444 de tabla de muestras, identificada por el tipo de caja "stbl", puede proporcionar ubicaciones (por ejemplo, ubicaciones con un archivo) para las muestras de medios en la pista, así como información de tiempo para las muestras. Usando la información proporcionada por la caja 444 de tabla de muestras, un dispositivo reproductor puede ubicar las muestras en el orden de tiempo correcto, determinar el tipo de muestra y/o determinar el tamaño, el contenedor y el desfase de una muestra dentro de un contenedor, entre otras cosas.

[0134] La caja 444 de tabla de muestras puede incluir una caja 446 de descripción de muestras, identificada por el tipo de caja "std". La caja 446 de descripción de muestras puede proporcionar información detallada sobre, por ejemplo, el tipo de codificación usado para una muestra y cualquier información de inicialización necesaria para ese tipo de codificación. La información almacenada en la caja de descripción de muestras puede ser específica para un tipo de pista que incluye las muestras. Por ejemplo, se puede usar un formato para la descripción de muestras cuando la pista es una pista de vídeo y se puede usar un formato diferente cuando la pista es una pista de indicaciones. Como ejemplo adicional, el formato de la descripción de muestras también puede variar según el formato de la pista de indicaciones.

[0135] La caja 446 de descripción de muestras puede incluir una o más entradas de muestras 448a a 448n. El tipo de entrada de muestras es una clase abstracta y, por lo tanto, típicamente la caja de descripción de muestras incluye una caja de entrada de muestras específica, como una entrada de muestras visuales para datos de vídeo o una entrada de muestras de audio para muestras de audio, entre otros ejemplos. Una caja de entrada de muestras puede almacenar los parámetros de una muestra en particular. Por ejemplo, para una muestra de vídeo, la caja de entrada de muestras puede incluir una anchura, altura, resolución horizontal, resolución vertical, recuento de tramas y/o profundidad para la muestra de vídeo, entre otras cosas. Como otro ejemplo, para una muestra de audio, la entrada de muestras puede incluir un recuento de canales, un diseño de canales y/o una frecuencia de muestreo, entre otras cosas.

[0136] En el ejemplo ilustrado, la primera entrada 448a de muestras incluye una caja 460 de información de esquema restringido. En tal ejemplo, la primera entrada 448a de muestras puede tener un código (por ejemplo, "resv") que indica un flujo restringido. Como se describe en el presente documento, una caja de información de esquema restringido, identificada por el tipo de caja "rinf", puede contener la información (por ejemplo, en una o más cajas) requerida tanto para comprender un esquema restringido aplicado a una muestra como los parámetros del esquema. Como se mencionó anteriormente, el autor de un archivo puede requerir ciertas acciones de un dispositivo reproductor para procesar cierto contenido de medios. En tales casos, el archivo puede incluir una caja de información de esquema restringido, que un dispositivo reproductor puede localizar y usar para determinar los requisitos para reproducir el contenido de medios asociado con la entrada de muestras. Los reproductores que no puedan reproducir el contenido también pueden usar la caja de información de esquema restringido para determinar que no pueden reproducir el contenido y, por lo tanto, no deben intentar procesar el contenido. La caja de información de esquema restringido típicamente incluye un tipo de entrada de muestras original, es decir, el tipo de entrada de muestras antes de cualquier transformación descrita por la caja de información de esquema restringido.

[0137] En varias implementaciones, como se señaló anteriormente, se pueden definir diferentes tipos de esquemas restringidos para diferentes tipos de contenido. El tipo de esquema restringido se puede especificar en una caja de tipo de esquema, identificada por el tipo de caja "schm". Ejemplos ilustrativos de tipos de esquemas restringidos incluyen un tipo de esquema restringido para disposiciones de vídeo estereoscópico (denotado como 'stvi'), un tipo de esquema restringido para vídeo de realidad virtual (denotado como 'vrvd'), un tipo de esquema restringido para vídeo omnidireccional proyectado (denotado como 'podv'), un tipo de esquema restringido para

vídeo proyectado equirrectangular (denotado como 'erpv'), un tipo de esquema restringido para vídeo proyectado equirrectangular o cubemap empaquetado (denotado como 'ercm'), un tipo de esquema restringido para vídeo omnidireccional de ojo de pez (denotado como 'fodv'), entre otros.

- 5 **[0138]** En el ejemplo ilustrativo mostrado en la FIG. 4, la primera entrada de muestras 448a incluye contenido de vídeo de realidad virtual estereoscópico empaquetado en tramas. Dos tipos de esquemas restringidos que son aplicables al contenido de vídeo de realidad virtual estereoscópico empaquetado en tramas incluyen el tipo de esquema restringido para disposiciones de vídeo estereoscópico ('stvi') y el tipo de esquema restringido para vídeo de realidad virtual ('vrvd'). La caja 460 de información de esquema restringido se agrega a la primera
- 10 entrada de muestras 448a que incluye el contenido de vídeo de realidad virtual estereoscópico empaquetado en tramas. La caja 462a de tipo de esquema indica el tipo de esquema restringido para contenido de realidad virtual, denotado como 'vrvd', y la caja 462b de tipo de esquema indica el tipo de esquema restringido para contenido de vídeo estereoscópico, denotado como 'stvi'.
- 15 **[0139]** La caja 460 de información de esquema restringido en el ejemplo ilustrado también incluye dos cajas 464a y 464b de información de esquema, ambas identificadas por el tipo de caja "schi". Las cajas 464a, 464b de información de esquema pueden almacenar información para los esquemas específicos para los que están relacionadas. Por ejemplo, la caja 464a de información de esquema es aplicable a la caja 462a de tipo de esquema restringido para el contenido de realidad virtual (el tipo de esquema 'vrvd'), e incluye parámetros para el
- 20 contenido de realidad virtual. Estos parámetros pueden incluir, por ejemplo, si el vídeo de realidad virtual es 2-D o 3-D, si el vídeo de realidad virtual está unido previamente o unido posteriormente y/o una correlación para el vídeo de realidad virtual. En varias implementaciones, se puede definir una caja de información de esquema para contenido de realidad virtual, específicamente para contener parámetros para el contenido de realidad virtual.
- 25 **[0140]** La caja 464b de información de esquema es aplicable a la caja 462b de tipo de esquema restringido para el contenido estereoscópico empaquetado de tramas (el tipo de esquema 'stvi'), e incluye los parámetros necesarios para el tipo de esquema restringido del contenido estereoscópico empaquetado de tramas. Por ejemplo, cuando se descodifican tramas de vídeo codificado en estéreo, las tramas descodificados pueden contener una representación de dos tramas constituyentes empaquetadas espacialmente que forman un par
- 30 estéreo (empaquetado de tramas) o pueden incluir solo una vista de un par estéreo (vistas izquierda y derecha en diferentes pistas). Las restricciones debidas al vídeo codificado en estéreo pueden incluirse en la caja 464b de información de esquema. Por ejemplo, la caja 464b de información de esquema es una caja contenedor, y una caja contenida en la caja 464b de información de esquema puede incluir una caja denominada 'StereoVideoBox'. En un ejemplo ilustrativo, para tramas de vídeo estereoscópico, los parámetros en la caja 464b de información de esquema (por ejemplo, en la StereoVideoBox contenida en la caja 464b de información de
- 35 esquema) pueden incluir información que indique que las tramas descodificadas contienen una representación de dos tramas constituyentes empaquetadas espacialmente que forman un par estéreo o contienen una de las dos vistas de un par estéreo. Por ejemplo, las dos imágenes constituyentes empaquetadas espacialmente pueden incluir dos imágenes de ojo de pez de la misma escena al mismo tiempo, pero desde diferentes vistas.
- 40 Las dos imágenes se pueden empaquetar en una trama de vídeo (por ejemplo, en una disposición de empaquetado lado a lado). En otro ejemplo, para imágenes omnidireccionales, los parámetros en la caja 464b de información de esquema pueden incluir información que indique que la imagen reconstruida contiene una representación de dos imágenes constituyentes empaquetadas espacialmente.
- 45 **[0141]** Ahora se describe un ejemplo ilustrativo de la Caja de vídeo estéreo. Como se indicó anteriormente, la caja de vídeo estéreo puede estar contenida en la caja 464b de información de esquema e incluye los datos necesarios para el tipo de esquema restringido 'stvi'. La caja de vídeo estéreo se puede definir de la siguiente manera:

Tipo de caja:	'stvi'
Contenedor:	SchemeInformationBox
Obligatorio:	Sí (cuando el tipo de esquema es 'stvi')
Cantidad:	Una

50

[0142] Un ejemplo de sintaxis y semántica para la caja de vídeo estéreo es el siguiente:

```
aligned(8) class StereoVideoBox extends FullBox('stvi', version
= 0, 0)
{
    template unsigned int(30) reserved = 0;
    unsigned int(2) single view allowed;
    unsigned int(32) stereo scheme;
    unsigned int(32) length;
    unsigned int (8) [length] stereo_indication_type;
```

```

    Box[] any_box; // opcional
}

```

single_view_allowed es un número entero. Un valor cero indica que el contenido solo se puede mostrar en pantallas estereoscópicas. Cuando (single_view_allowed & 1) es igual a 1, se permite mostrar la vista derecha en una pantalla monoscópica de vista única. Cuando (single_view_allowed & 2) es igual a 2, se permite mostrar la vista izquierda en una pantalla monoscópica de vista única. stereo_scheme es un número entero que indica el esquema de disposición estéreo utilizado y el tipo de indicación estéreo de acuerdo con el esquema utilizado. Se especifican los siguientes valores para stereo_scheme:

1: el esquema de empaquetado de tramas como se especifica en el mensaje de información complementaria de mejora de la disposición de empaquetado de tramas de ISO/IEC 14496-10 [ISO/IEC 14496-10]

2: el esquema de tipo de disposición como se especifica en el Anexo L de ISO/IEC 13818-2 [ISO/IEC 13818-2:2000/Amd.4]

3: el esquema estéreo como se especifica en ISO/IEC 23000-11 para los servicios mixtos 2D/3D y compatibles con trama/servicio.

[0143] Otros valores de stereo_scheme están reservados. length indica el número de bytes para el campo stereo_indication_type. stereo_indication_type indica el tipo de disposición estéreo de acuerdo con el esquema de indicación estéreo utilizado. La sintaxis y la semántica de stereo_indication_type dependen del valor de stereo_scheme. La sintaxis y la semántica de stereo_indication_type para los siguientes valores de stereo_scheme se especifican de la siguiente manera:

stereo_scheme igual a 1: El valor de length será 4 y stereo_indication_type será unsigned int (32) que contiene el valor frame_packing_arrangement_type de la Tabla D-8 de ISO/IEC 14496-10 [ISO/IEC 14496-10] ('Definición de frame_packing_arrangement_type').

stereo_scheme igual a 2: El valor de la length será 4 y stereo_indication_type será unsigned int (32) que contiene el valor de tipo de la Tabla L-1 de ISO/IEC 13818-2 [ISO/IEC 13818-2:2000/Amd.4] ('Definición de arrangement_type').

stereo_scheme igual a 3: El valor de length será 2 y stereo_indication_type contendrá dos elementos de sintaxis de unsigned int (8). El primer elemento de sintaxis deberá contener el tipo de composición estereoscópica de la Tabla 4 de ISO/IEC 23000-11:2009. El bit menos significativo del segundo elemento de sintaxis contendrá el valor de is_left_first como se especifica en 8.4.3 de ISO/IEC 23000-11:2009, mientras que los otros bits están reservados y se pondrán a 0.

[0144] Lo siguiente se aplica cuando se utiliza StereoVideoBox:

- En TrackHeaderBox

- ancho y alto especifican el tamaño de presentación visual de una sola vista después de desembalar.

- En SampleDescriptionBox

- frame_count será 1, porque el decodificador genera físicamente una sola trama. En otras palabras, frame_count no documenta las tramas constituyentes incluidas en una imagen empaquetada.

- el ancho y la altura documentan los recuentos de píxeles de una imagen empaquetada en tramas (y no los recuentos de píxeles de una vista única dentro de una imagen empaquetada en tramas).

- PixelAspectRatioBox documenta la relación de aspecto de píxeles de cada vista cuando la vista se muestra en una pantalla monoscópica de una sola vista. Por ejemplo, en muchas disposiciones de empaquetado espacial de tramas, la caja Proporción de aspecto de píxel indica una proporción de aspecto de píxel de 2:1 o 1:2, ya que la resolución espacial de una vista de vídeo empaquetado en tramas típicamente se reduce a la mitad a lo largo de un eje de coordenadas en comparación con ese del vídeo de vista única del mismo formato.

[0145] Al generar la caja 442 de información de medios con la caja 460 de información de esquema restringido para incluir tanto la caja 462a de tipo de esquema como la caja 462b de tipo de esquema, así como la caja 464b de información de esquema como la caja 464b de información de esquema, ambos tipos de esquemas restringidos (los tipos de esquema 'vrvd' y 'stvi') se pueden aplicar simultáneamente al contenido de vídeo de realidad virtual estereoscópico empaquetado en tramas asociado con la primera entrada de muestras 448a.

[0146] Existen muchos otros ejemplos en los que se pueden aplicar múltiples esquemas restringidos al mismo contenido de medios asociado con una entrada de muestras. Por ejemplo, el contenido de vídeo de realidad virtual (o vídeo de 360 grados) puede incluir contenido de vídeo esférico proyectado en una representación equirrectangular. Como otro ejemplo, el contenido de vídeo de realidad virtual (o vídeo de 360 grados) puede incluir contenido de vídeo esférico proyectado en una geometría de mapa de cubos, una geometría de pirámide cuadrada truncada u otra proyección. Se pueden aplicar múltiples esquemas restringidos a dicho contenido. A continuación, se describen ejemplos de esquemas restringidos que se pueden incluir en un archivo de medios y que son aplicables al contenido de medios asociado con una entrada de muestras en el archivo de medios.

[0147] Otro ejemplo de un tipo de esquema restringido es para vídeo omnidireccional proyectado (con un tipo de esquema denotado como 'podv'). Por ejemplo, el uso del esquema de vídeo omnidireccional proyectado para el tipo de entrada de muestras de vídeo restringido 'resv' indica que las imágenes descodificadas son imágenes empaquetadas que contienen contenido monoscópico o estereoscópico. El uso del esquema de vídeo omnidireccional proyectado se indica mediante el `schema_type` igual a 'podv' (vídeo omnidireccional proyectado) dentro de una caja de tipo de esquema (`SchemaTypeBox`) en una caja de información de esquema restringido (`RestrictedSchemeInfoBox`). El formato de las imágenes monoscópicas proyectadas se indica con una `ProjectedOmniVideoBox` contenida dentro de la caja de información de esquema (`SchemeInformationBox`) en la caja de información de esquema restringido. En algunos casos, una y solo una `ProjectedOmniVideoBox` estará presente en la `SchemeInformationBox` cuando el tipo de esquema es 'podv'. En algunos casos, el tipo de esquema 'podv' se define como un tipo de esquema abierto para vídeo omnidireccional proyectado.

[0148] Se proporciona otro tipo de esquema restringido para vídeo proyectado equirrectangular (con un tipo de esquema denotado como 'erpv'). En algunos casos, el tipo de esquema 'erpv' se define como un tipo de esquema cerrado para vídeo omnidireccional proyectado. Cuando `schema_type` es igual a 'erpv' en una caja de tipo de esquema (también denominada caja de tipo de esquema compatible) en una caja de información de esquema restringido (`RestrictedSchemeInfoBox`), la pista correspondiente se ajusta a las restricciones de `schema_type` igual a 'podv' (como se define en los documentos de normas ISO/IEC 14496-15) con varias restricciones adicionales. Por ejemplo, una caja de formato de proyección (`ProjectionFormatBox`) dentro de una caja de vídeo omnidireccional proyectado (`ProjectedOmniVideoBox`) indicará la proyección equirrectangular. Además, cuando está presente una caja de empaquetado regional (`RegionWisePackingBox`), se aplican todas las siguientes restricciones:

[0149] El valor de `NumRegions` será igual a $\text{HorDiv1} * \text{VerDiv1}$.

[0150] Para cada valor de `i` en el rango de 0 a `NumRegions - 1`, inclusive, se aplica lo siguiente:

El valor de `PackingType[i]` será igual a 0.

[0151] El valor de `TransformType[i]` será igual a 0.

[0152] El valor de `PackedRegWidth[i]` será igual a `ProjRegWidth[i]`.

[0153] El valor de `PackedRegHeight[i]` será igual a `ProjRegHeight[i]`.

[0154] Otras restricciones asociadas con el tipo de esquema restringido de vídeo proyectado equirrectangular 'erpv' incluyen que la versión de `ProjectionFormatBox`, `StereoVideoBox` (cuando está presente), `RegionWisePackingBox` (cuando está presente), `RotationBox` (cuando está presente) y `CoverageInformationBox` (cuando está presente) serán igual a 0. Además, la caja de información de esquema (`SchemeInformationBox`) no debe contener directa o indirectamente ninguna caja que no sea `ProjectedOmniVideoBox`, `ProjectionFormatBox`, `StereoVideoBox`, `RegionWisePackingBox`, `RotationBox` y `CoverageInformationBox`.

[0155] La FIG. 5 es un diagrama de flujo que ilustra un ejemplo de un proceso 500 para procesar datos de vídeo usando las técnicas descritas en el presente documento. En el bloque 502, el proceso 500 incluye obtener los datos de vídeo. En un ejemplo ilustrativo, los datos de vídeo pueden incluir una parte (por ejemplo, una o más muestras) de un flujo de bits de vídeo codificado que se asociará con una entrada de muestras de un archivo de formato de medios. En algunos ejemplos, el flujo de bits de vídeo codificado se puede cifrar. En otro ejemplo ilustrativo, los datos de vídeo pueden incluir una porción de vídeo (por ejemplo, capturado por una o más cámaras) que se asociará con una entrada de muestras de un archivo de formato de medios. En tal ejemplo, los datos de vídeo pueden codificarse y, en algunos casos, cifrarse mediante el proceso 500 antes de que se genere el archivo de formato de medios.

[0156] En el bloque 504, el proceso 500 incluye generar un archivo de formato de medios que incluye los datos de vídeo y una caja de información de esquema restringido asociado con los datos de vídeo. En un ejemplo ilustrativo, el archivo de formato de medios se genera usando un formato de archivo de medios de base ISO (ISOBMFF) o usando un formato de archivo de medios derivado del formato de archivo de medios de base ISO.

La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo. Por ejemplo, cada caja de tipo de esquema de la pluralidad de cajas de tipo de esquema puede identificar un tipo de esquema restringido diferente requerido para procesar los datos de vídeo. En algunos casos, los datos de vídeo están asociados con una entrada de muestras del archivo de formato de medios. En tales casos, el archivo de formato de medios puede especificar múltiples esquemas restringidos (usando la pluralidad de cajas de tipo de esquema y, en algunos casos, múltiples cajas de información de esquema) que son aplicables a datos de vídeo asociados con una entrada de muestras del archivo de formato de medios. En algunos ejemplos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica una versión de un tipo de esquema restringido identificado por la caja de tipo de esquema. Por ejemplo, el valor de `scheme_version` usado anteriormente puede tener un valor particular (por ejemplo, 0, 1, 2, etc.) que indica una versión de un tipo de esquema restringido particular. Las diferentes versiones de un tipo de esquema restringido particular pueden incluir variantes del esquema restringido indicado por el tipo de esquema. En algunos ejemplos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema incluye un identificador que identifica una fuente de datos necesaria para implementar un tipo de esquema restringido identificado por la caja de tipo de esquema. Por ejemplo, el identificador puede incluir un localizador uniforme de recursos (URL) u otro identificador de fuente adecuado que pueda proporcionar acceso a una fuente para obtener datos.

[0157] Los tipos de esquemas restringidos identificados por la pluralidad de cajas de tipo de esquema pueden incluir cualquier tipo de esquema restringido adecuado. Ejemplos ilustrativos de tipos de esquemas restringidos incluyen un tipo de esquema restringido para disposiciones de vídeo estereoscópico (denotado como 'stvi'), un tipo de esquema restringido para vídeo de realidad virtual (denotado como 'vrvd'), un tipo de esquema restringido para vídeo omnidireccional proyectado (denotado como 'podv'), un tipo de esquema restringido para vídeo proyectado equirrectangular (denotado como 'erpv'), un tipo de esquema restringido para vídeo proyectado equirrectangular o cubemap empaquetado (denotado como 'ercm'), un tipo de esquema restringido para vídeo omnidireccional de ojo de pez (denotado como 'fodv'), entre otros.

[0158] En algunos ejemplos, la caja de información de esquema restringido incluye, además, una pluralidad de cajas de información de esquema. Una caja de información de esquema de la pluralidad de cajas de información de esquema incluye los datos necesarios para un tipo de esquema restringido identificado por una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema. Utilizando un ejemplo anterior, la caja de información de esquema puede incluir una caja de vídeo estéreo que incluye los datos necesarios para el tipo de esquema restringido 'stvi'.

[0159] Una vez generada, el archivo de formato de medios se puede utilizar de cualquier manera adecuada. Por ejemplo, el proceso 500 puede incluir la transmisión del archivo de formato de medios. Por ejemplo, el archivo de formato de medios puede transmitirse a un dispositivo de procesamiento de archivos (por ejemplo, un reproductor de medios, un analizador de archivos, un dispositivo de decodificación, un dispositivo de almacenamiento, un dispositivo intermediario como un servidor y/u otro dispositivo adecuado). En otro ejemplo, el proceso 500 puede incluir almacenar el archivo de formato de medios.

[0160] En algunos ejemplos, el proceso 500 puede ser realizado por un dispositivo informático o un aparato, tal como un dispositivo de generación de archivos de formato de medios, el dispositivo de codificación 104 (que puede configurarse para generar archivos de formato de medios) o cualquier otro dispositivo informático. En algunos casos, el dispositivo o aparato informático puede incluir un procesador, microprocesador, microordenador u otro componente de un dispositivo que esté configurado para llevar a cabo los pasos del proceso 500. En algunos ejemplos, el dispositivo o aparato informático puede incluir una cámara configurada para capturar datos de vídeo (por ejemplo, una secuencia de vídeo) que incluye tramas de vídeo. Por ejemplo, el dispositivo informático puede incluir un dispositivo de cámara, que puede incluir o no un códec de vídeo. Como otro ejemplo, el dispositivo informático puede incluir un dispositivo móvil con una cámara (por ejemplo, un dispositivo de cámara como una cámara digital, una cámara IP o similar, un teléfono móvil o tableta que incluya una cámara u otro tipo de dispositivo con una cámara). En algunos casos, el dispositivo informático puede incluir una pantalla para mostrar imágenes. En algunos ejemplos, una cámara u otro dispositivo de captura que captura los datos de vídeo está separado del dispositivo informático, en cuyo caso el dispositivo informático recibe los datos de vídeo capturados. El dispositivo informático puede incluir, además, una interfaz de red configurada para comunicar los datos de vídeo. La interfaz de red puede configurarse para comunicar datos basados en el Protocolo de Internet (IP) u otros datos de red.

[0161] La FIG. 6 es un diagrama de flujo que ilustra un ejemplo de un proceso 600 para procesar uno o más archivos de formato de medios usando las técnicas descritas en el presente documento. En el bloque 602, el proceso 600 incluye obtener un archivo de formato de medios que incluye datos de vídeo y una caja de información de esquema restringido asociada con los datos de vídeo. En un ejemplo ilustrativo, el archivo de formato de medios se genera usando un formato de archivo de medios de base ISO (ISOBMFF) o usando un formato de archivo de medios derivado del formato de archivo de medios de base ISO. La caja de información de esquema restringido incluye una pluralidad de cajas de tipo de esquema. La pluralidad de cajas de tipo de

esquema identifica una pluralidad de tipos de esquema restringido necesarios para procesar los datos de vídeo. Por ejemplo, cada caja de tipo de esquema de la pluralidad de cajas de tipo de esquema puede identificar un tipo de esquema restringido diferente requerido para procesar los datos de vídeo. En algunos casos, los datos de vídeo están asociados con una entrada de muestras del archivo de formato de medios. En tales casos, el archivo de formato de medios puede especificar múltiples esquemas restringidos (usando la pluralidad de cajas de tipo de esquema y, en algunos casos, múltiples cajas de información de esquema) que son aplicables a datos de vídeo asociados con una entrada de muestras del archivo de formato de medios. En algunos ejemplos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica una versión de un tipo de esquema restringido identificado por la caja de tipo de esquema. En algunos ejemplos, una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema incluye un identificador que identifica una fuente de datos necesaria para implementar un tipo de esquema restringido identificado por la caja de tipo de esquema. Por ejemplo, el identificador puede incluir un localizador uniforme de recursos (URL) u otro identificador de fuente adecuado que pueda proporcionar acceso a una fuente para obtener datos.

[0162] Los tipos de esquemas restringidos identificados por la pluralidad de cajas de tipo de esquema pueden incluir cualquier tipo de esquema restringido adecuado. Ejemplos ilustrativos de tipos de esquemas restringidos incluyen un tipo de esquema restringido para disposiciones de vídeo estereoscópico (denotado como 'stvi'), un tipo de esquema restringido para vídeo de realidad virtual (denotado como 'vrvd'), un tipo de esquema restringido para vídeo omnidireccional proyectado (denotado como 'podv'), un tipo de esquema restringido para vídeo proyectado equirrectangular (denotado como 'erpv'), un tipo de esquema restringido para vídeo proyectado equirrectangular o cubemap empaquetado (denotado como 'ercm'), un tipo de esquema restringido para vídeo omnidireccional de ojo de pez (denotado como 'fodv'), entre otros.

[0163] En algunos ejemplos, la caja de información de esquema restringido incluye, además, una pluralidad de cajas de información de esquema. Una caja de información de esquema de la pluralidad de cajas de información de esquema incluye los datos necesarios para un tipo de esquema restringido identificado por una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema. Utilizando un ejemplo anterior, la caja de información de esquema puede incluir una caja de vídeo estéreo que incluye los datos necesarios para el tipo de esquema restringido 'stvi'.

[0164] En el bloque 604, el proceso 600 incluye analizar el archivo de formato de medios. Por ejemplo, cada una de las cajas incluidas en el archivo de formato de medios se puede analizar desenvolviendo las cajas y obteniendo el contenido dentro de las cajas. En el bloque 606, el proceso 600 incluye procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquemas restringidos identificados por la pluralidad de cajas de tipo de esquema.

[0165] Los datos de vídeo y el contenido analizados del archivo de formato de medios se pueden procesar según sea necesario. Por ejemplo, el proceso 600 puede descodificar los datos de vídeo (por ejemplo, usando un descodificador). El proceso 600 puede incluir, además, la representación y muestra de los datos de vídeo descodificados usando la información relacionada con los esquemas de restricción incluidos en el archivo de formato de medios.

[0166] En algunos ejemplos, el proceso 600 puede ser realizado por un dispositivo o un aparato informático, tal como un dispositivo de procesamiento de archivos de formato de medios, el dispositivo de descodificación 112 (que se puede configurar para analizar archivos de formato de medios, por ejemplo, realizando transformaciones inversas de la archivos de formato de medios), o cualquier otro dispositivo informático. En algunos casos, el dispositivo o aparato informático puede incluir un procesador, microprocesador, microordenador u otro componente de un dispositivo que esté configurado para llevar a cabo los pasos del proceso 600. En algunos ejemplos, el dispositivo o aparato informático puede incluir una cámara configurada para capturar datos de vídeo (por ejemplo, una secuencia de vídeo) que incluye tramas de vídeo. Por ejemplo, el dispositivo informático puede incluir un dispositivo de cámara, que puede incluir o no un códec de vídeo. Como otro ejemplo, el dispositivo informático puede incluir un dispositivo móvil con una cámara (por ejemplo, un dispositivo de cámara como una cámara digital, una cámara IP o similar, un teléfono móvil o tableta que incluya una cámara u otro tipo de dispositivo con una cámara). En algunos casos, el dispositivo informático puede incluir una pantalla para mostrar imágenes. En algunos ejemplos, una cámara u otro dispositivo de captura que captura los datos de vídeo está separado del dispositivo informático, en cuyo caso el dispositivo informático recibe los datos de vídeo capturados. El dispositivo informático puede incluir, además, una interfaz de red configurada para comunicar los datos de vídeo. La interfaz de red puede configurarse para comunicar datos basados en el Protocolo de Internet (IP) u otros datos de red.

[0167] Los procesos 500 y 600 se ilustran como un diagrama de flujo lógico, cuyo funcionamiento representa una secuencia de operaciones que pueden implementarse en hardware, instrucciones de ordenador o una combinación de los mismos. En el contexto de las instrucciones de ordenador, las operaciones representan instrucciones ejecutables por ordenador almacenadas en uno o más medios de almacenamiento legibles por ordenador que, cuando son ejecutadas por uno o más procesadores, realizan las operaciones enumeradas. En general, las instrucciones ejecutables por ordenador incluyen rutinas, programas, objetos, componentes,

estructuras de datos y similares que realizan funciones particulares o implementan tipos de datos particulares. El orden en el que se describen las operaciones no debe interpretarse como una limitación, y cualquier número de las operaciones descritas se puede combinar en cualquier orden y/o en paralelo para implementar los procesos.

[0168] Además, los procesos 500 y 600 pueden realizarse bajo el control de uno o más sistemas informáticos configurados con instrucciones ejecutables y pueden implementarse como código (por ejemplo, instrucciones ejecutables, uno o más programas informáticos, o una o más aplicaciones) ejecutándose colectivamente en uno o más procesadores, mediante hardware o combinaciones de los mismos. Como se indicó anteriormente, el código puede almacenarse en un medio de almacenamiento legible por ordenador o legible por máquina, por ejemplo, en forma de un programa informático que comprende una pluralidad de instrucciones ejecutables por uno o más procesadores. El medio de almacenamiento legible por ordenador o legible por máquina puede no ser transitorio.

[0169] Las técnicas y sistemas analizados en el presente documento pueden implementarse en un sistema de codificación y decodificación de vídeo de ejemplo (por ejemplo, el sistema 100). En algunos ejemplos, un sistema incluye un dispositivo de origen que proporciona datos de vídeo codificados, a decodificar en un momento posterior mediante un dispositivo de destino. En particular, el dispositivo de origen proporciona los datos de vídeo al dispositivo de destino a través de un medio legible por ordenador. El dispositivo de origen y el dispositivo de destino pueden comprender cualquiera de una amplia gama de dispositivos, incluyendo ordenadores de escritorio, notebooks (es decir, ordenadores portátiles), tabletas electrónicas, decodificadores, teléfonos tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión continua de vídeo o similares. En algunos casos, el dispositivo de origen y el dispositivo de destino pueden estar equipados para la comunicación inalámbrica.

[0170] El dispositivo de destino puede recibir los datos de vídeo codificados a decodificar mediante el medio legible por ordenador. El medio legible por ordenador puede comprender cualquier tipo de medio o dispositivo capaz de transportar los datos de vídeo codificados desde el dispositivo de origen al dispositivo de destino. En un ejemplo, el medio legible por ordenador puede comprender un medio de comunicación para habilitar el dispositivo de origen para transmitir datos de vídeo codificados directamente al dispositivo de destino en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo de destino. El medio de comunicación puede comprender cualquier medio de comunicación inalámbrico o por cable, tal como un espectro de radiofrecuencia (RF) o una o más líneas físicas de transmisión. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global, tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen hasta el dispositivo de destino.

[0171] En algunos ejemplos, pueden proporcionarse datos codificados desde la interfaz de salida hasta un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento mediante una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera de una variedad de medios de almacenamiento de datos, distribuidos o de acceso local, tales como un disco duro, discos Blu-ray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil, o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de vídeo codificados. En un ejemplo adicional, el dispositivo de almacenamiento puede corresponder a un servidor de archivos o a otro dispositivo de almacenamiento intermedio que pueda almacenar el vídeo codificado generado por el dispositivo de origen. El dispositivo de destino puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento mediante transmisión continua o descarga. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino. Los servidores de archivos de ejemplo incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión Wi-Fi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambas cosas que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0172] Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de cualquiera de una variedad de aplicaciones de multimedia, tales como radiodifusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones continua de vídeo por internet, tales como la transmisión continua adaptativa dinámica por HTTP (DASH), vídeo digital que se codifica en un medio de almacenamiento de datos, decodificación de vídeo digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema puede estar configurado para prestar soporte a una

transmisión de vídeo unidireccional o bidireccional, a fin de prestar soporte a aplicaciones tales como la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0173] En un ejemplo, el dispositivo de origen incluye una fuente de vídeo, un codificador de vídeo y una interfaz de salida. El dispositivo de destino puede incluir una interfaz de entrada, un descodificador de vídeo y un dispositivo de visualización. El codificador de vídeo del dispositivo de origen puede configurarse para aplicar las técnicas divulgadas en el presente documento. En otros ejemplos, un dispositivo de origen y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen puede recibir datos de vídeo desde una fuente de vídeo externa, tal como una cámara externa. Del mismo modo, el dispositivo de destino puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

[0174] El sistema de ejemplo anterior es simplemente un ejemplo. Las técnicas para el procesamiento de datos de vídeo en paralelo pueden realizarse mediante cualquier dispositivo de codificación y/o de descodificación de vídeo digital. Aunque, en general, las técnicas de esta divulgación se realizan mediante un dispositivo de codificación de vídeo, las técnicas también se pueden realizar mediante un codificador/descodificador de vídeo, denominado típicamente "CÓDEC". Además, las técnicas de esta divulgación también se pueden realizar mediante un preprocesador de vídeo. El dispositivo de origen y el dispositivo de destino son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen genera datos de vídeo codificados para su transmisión al dispositivo de destino. En algunos ejemplos, los dispositivos de origen y destino pueden funcionar de manera esencialmente simétrica, de modo que cada uno de los dispositivos incluye componentes de codificación y de descodificación de vídeo. De ahí que los sistemas de ejemplo puedan prestar soporte a la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo o la videotelefonía.

[0175] La fuente de vídeo puede incluir un dispositivo de captura de vídeo, tal como una videocámara, un archivo de vídeo que contiene vídeo capturado previamente y/o una interfaz de suministro de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. Como otra alternativa, la fuente de vídeo puede generar datos basados en gráficos de ordenador como el vídeo de origen, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente de vídeo es una videocámara, el dispositivo de origen y el dispositivo de destino pueden formar los denominados teléfonos con cámara o videoteléfonos. Sin embargo, como se menciona anteriormente, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o por cable. En cada caso, el codificador de vídeo puede codificar el vídeo capturado, precapturado o generado por ordenador. A continuación, la información de vídeo codificada puede ser proporcionada por la interfaz de salida a un medio legible por ordenador.

[0176] Como se ha indicado, el medio legible por ordenador puede incluir medios transitorios, tales como una radiodifusión inalámbrica o una transmisión de red por cable, o medios de almacenamiento (es decir, medios de almacenamiento no transitorios), tales como un disco duro, una unidad de memoria flash, un disco compacto, un disco de vídeo digital, un disco Blu-ray u otros medios legibles por ordenador. En algunos ejemplos, un servidor de red (no mostrado) puede recibir datos de vídeo codificados desde el dispositivo de origen y proporcionar los datos de vídeo codificados al dispositivo de destino, por ejemplo, por medio de una transmisión en red. De forma similar, un dispositivo informático de una unidad de producción de medios, tal como una unidad de grabación de discos, puede recibir datos de vídeo codificados desde el dispositivo de origen y producir un disco que contiene los datos de vídeo codificados. Por lo tanto, se puede entender que el medio legible por ordenador incluye uno o más medios legibles por ordenador de diversas formas, en varios ejemplos.

[0177] La interfaz de entrada del dispositivo de destino recibe información del medio legible por ordenador. La información del medio legible por ordenador puede incluir información de sintaxis definida por el codificador de vídeo, que también es usada por el descodificador de vídeo, que incluye elementos de sintaxis que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, grupos de imágenes (GOP). Un dispositivo de visualización muestra los datos de vídeo descodificados a un usuario, y puede comprender cualquiera de una variedad de dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización. Se han descrito diversos modos de realización de la invención.

[0178] Los detalles específicos del dispositivo de codificación 104 y del dispositivo de descodificación 112 se muestran en la FIG. 7 y la FIG. 8, respectivamente. La FIG. 7 es un diagrama de bloques que ilustra un dispositivo de codificación 104 de ejemplo que puede implementar una o más de las técnicas descritas en esta divulgación. El dispositivo de codificación 104 puede, por ejemplo, generar las estructuras de sintaxis descritas en el presente documento (por ejemplo, las estructuras de sintaxis de un VPS, SPS, PPS u otros elementos de sintaxis). El dispositivo de codificación 104 puede realizar codificación de intrapredicción e interpredicción de bloques de vídeo dentro de fragmentos de vídeo. Como se describe previamente, la intracodificación se basa, al menos en parte, en la predicción espacial para reducir o eliminar la redundancia espacial en una imagen o trama

de vídeo dada. La intercodificación se basa, al menos en parte, en la predicción temporal para reducir o eliminar la redundancia temporal en las tramas adyacentes o circundantes de una secuencia de vídeo. El intramodo (modo I) se puede referir a cualquiera de varios modos de compresión de base espacial. Los intermodos, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), se pueden referir a cualquiera de varios modos de compresión de base temporal.

[0179] El dispositivo de codificación 104 incluye una unidad de división 35, una unidad de procesamiento de predicción 41, una unidad de filtro 63, una memoria de imágenes 64, un sumador 50, una unidad de procesamiento de transformación 52, una unidad de cuantificación 54 y una unidad de codificación por entropía 56. La unidad de procesamiento de predicción 41 incluye una unidad de estimación de movimiento 42, una unidad de compensación de movimiento 44 y una unidad de procesamiento de intrapredicción 46. Para la reconstrucción de bloques de vídeo, el dispositivo de codificación 104 incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformación inversa 60 y un sumador 62. La unidad de filtro 63 pretende representar uno o más filtros de bucle tales como un filtro de desbloqueo, un filtro de bucle adaptativo (ALF) y un filtro de desviación adaptativa de muestras (SAO). Aunque la unidad de filtro 63 se muestra en la FIG. 7 siendo un filtro de bucle, en otras configuraciones la unidad de filtro 63 se puede implementar como un filtro de bucle posterior. Un dispositivo de procesamiento posterior 57 puede realizar un procesamiento adicional en los datos de vídeo codificados generados por el dispositivo de codificación 104. Las técnicas de esta divulgación pueden implementarse en algunos casos mediante el dispositivo de codificación 104. En otros casos, sin embargo, una o más de las técnicas de esta divulgación pueden implementarse mediante el dispositivo de procesamiento posterior 57.

[0180] Como se muestra en la FIG. 7, el dispositivo de codificación 104 recibe datos de vídeo, y la unidad de división 35 divide los datos en bloques de vídeo. La división también puede incluir la división en fragmentos, segmentos de fragmento, mosaicos u otras unidades mayores, así como la división en bloques de vídeo, por ejemplo, de acuerdo con una estructura de árbol cuaternario de LCU y CU. El dispositivo de codificación 104 ilustra, en general, los componentes que codifican bloques de vídeo dentro de un fragmento de vídeo que se va a codificar. El fragmento se puede dividir en múltiples bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo denominados mosaicos). La unidad de procesamiento de predicción 41 puede seleccionar uno de una pluralidad de posibles modos de codificación, tal como uno de una pluralidad de modos de codificación de intrapredicción, o uno de una pluralidad de modos de codificación de interpredicción, para el bloque de vídeo actual basándose en resultados de error (por ejemplo, la velocidad de codificación y el nivel de distorsión, o similares). La unidad de procesamiento de predicción 41 puede proporcionar el bloque intracodificado o intercodificado resultante al sumador 50 para generar datos de bloque residuales y al sumador 62 para reconstruir el bloque codificado para su uso como imagen de referencia.

[0181] La unidad de procesamiento de intrapredicción 46, situada en la unidad de procesamiento de predicción 41, puede realizar la codificación de intrapredicción del bloque de vídeo actual con respecto a uno o más bloques vecinos en la misma trama o fragmento que el bloque actual que se va a codificar, para proporcionar compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 dentro de la unidad de procesamiento de predicción 41 realizan la codificación interpredictiva del bloque de vídeo actual en relación con uno o más bloques predictivos en una o más imágenes de referencia, para proporcionar compresión temporal.

[0182] La unidad de estimación de movimiento 42 se puede configurar para determinar el modo de interpredicción para un fragmento de vídeo de acuerdo con un patrón predeterminado para una secuencia de vídeo. El patrón predeterminado puede designar fragmentos de vídeo de la secuencia como fragmentos P, fragmentos B o fragmentos GPB. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado para propósitos conceptuales. La estimación de movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una unidad de predicción (PU) de un bloque de vídeo dentro de una trama o imagen de vídeo actual con respecto a un bloque predictivo dentro de una imagen de referencia.

[0183] Un bloque predictivo es un bloque que se encuentra que coincide estrechamente con la PU del bloque de vídeo que se va a codificar en términos de diferencia de píxel, que se puede determinar mediante una suma de diferencia absoluta (SAD), una suma de diferencia al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el dispositivo de codificación 104 puede calcular los valores para las posiciones fraccionarias de píxel de imágenes de referencia almacenadas en la memoria de imágenes 64. Por ejemplo, el dispositivo de codificación 104 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativa a las posiciones de píxel completo y a las posiciones de píxel fraccionario, y emitir un vector de movimiento con una precisión de píxel fraccionario.

[0184] La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento intercodificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse entre una primera lista de imágenes de referencia (Lista 0) o una segunda lista de imágenes de referencia (Lista 1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en la memoria de imágenes 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

[0185] La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo en base al vector de movimiento determinado mediante la estimación de movimiento, realizando posiblemente interpolaciones con una precisión de subpíxel. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una lista de imágenes de referencia. El dispositivo de codificación 104 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, formando valores de diferencia de píxel. Los valores de diferencia de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencia tanto de luma como de croma. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de compensación de movimiento 44 también puede generar elementos de sintaxis asociados a los bloques de vídeo y al fragmento de vídeo para su uso mediante el dispositivo de descodificación 112 en la descodificación de los bloques de vídeo del fragmento de vídeo.

[0186] La unidad de procesamiento de intrapredicción 46 puede intrapredecir un bloque actual, como una alternativa a la interpredicción realizada por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se describe anteriormente. En particular, la unidad de procesamiento de intrapredicción 46 puede determinar un modo de intrapredicción para su uso en la codificación de un bloque actual. En algunos ejemplos, la unidad de procesamiento de intrapredicción 46 puede codificar un bloque actual usando diversos modos de intrapredicción, por ejemplo, durante pasadas de codificación independientes, y la unidad de procesamiento de intrapredicción 46 puede seleccionar un modo de intrapredicción adecuado para su uso a partir de los modos probados. Por ejemplo, la unidad de procesamiento de intrapredicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para los diversos modos de intrapredicción sometidos a prueba, y puede seleccionar el modo de intrapredicción que tenga las mejores características de velocidad-distorsión entre los modos sometidos a prueba. El análisis de velocidad-distorsión determina, en general, una cantidad de distorsión (o de errores) entre un bloque codificado y un bloque original, no codificado, que se codificó para producir el bloque codificado, así como una velocidad de transferencia de bits (es decir, un número de bits) usada para producir el bloque codificado. La unidad de procesamiento de intrapredicción 46 puede calcular relaciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modo de intrapredicción presenta el mejor valor de velocidad-distorsión para el bloque.

[0187] En cualquier caso, tras seleccionar un modo de intrapredicción para un bloque, la unidad de procesamiento de intrapredicción 46 puede proporcionar información indicativa del modo de intrapredicción seleccionado para el bloque a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica el modo de intrapredicción seleccionado. El dispositivo de codificación 104 puede incluir en la configuración de flujo de bits transmitido definiciones de datos de contextos de codificación para varios bloques, así como indicaciones de un modo de predicción intra más probable, una tabla de índice de modo predicción intra y una tabla de índice de modo predicción intra modificado para utilizar para cada uno de los contextos. Los datos de configuración de flujo de bits pueden incluir una pluralidad de tablas de índices de modo de intrapredicción y una pluralidad de tablas de índices de modo de intrapredicción modificadas (también denominadas tablas de correlación de palabras de código).

[0188] Después de que la unidad de procesamiento de predicción 41 genere el bloque predictivo para el bloque de vídeo actual, ya sea por medio de interpredicción o intrapredicción, el dispositivo de codificación 104 forma un bloque de vídeo residual restando el bloque predictivo al bloque de vídeo actual. Los datos de vídeo residuales del bloque residual se pueden incluir en una o más TU y aplicarse a la unidad de procesamiento de transformación 52. La unidad de procesamiento de transformación 52 transforma los datos de vídeo residuales en coeficientes de transformación residuales usando una transformación, tal como una transformación de coseno discreta (DCT) o una transformación conceptualmente similar. La unidad de procesamiento de transformación 52 puede convertir los datos de vídeo residuales de un dominio de píxel a un dominio de transformación, tal como un dominio de frecuencia.

[0189] La unidad de procesamiento de transformación 52 puede enviar los coeficientes de transformación resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformación para reducir aún más la tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos, o con la totalidad, de los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede realizar, a

continuación, una exploración de la matriz que incluye los coeficientes de transformación cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede realizar la exploración.

[0190] Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformación cuantificados. Por ejemplo, la unidad de codificación por entropía 56 puede llevar a cabo la codificación de longitud variable adaptativa al contexto (CAVLC), la codificación aritmética binaria adaptativa el contexto (CABAC), la codificación aritmética binaria adaptativa el contexto basada en sintaxis (SBAC), la codificación por entropía por división en intervalos de probabilidad (PIPE) u otra técnica de codificación por entropía. Seguidamente a la codificación por entropía mediante la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir al dispositivo de decodificación 112, o archivarse para su posterior transmisión o recuperación mediante el dispositivo de decodificación 112. La unidad de codificación por entropía 56 también puede codificar por entropía los vectores de movimiento y los otros elementos de sintaxis para el fragmento de vídeo actual que se esté codificando.

[0191] La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformación inversa 60 aplican una cuantificación inversa y una transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxel para su uso posterior como bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia de una lista de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxeles fraccionarios, para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque predictivo con compensación de movimiento generado por la unidad de compensación de movimiento 44 para generar un bloque de referencia para su almacenamiento en la memoria de imágenes 64. El bloque de referencia puede ser usado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 como bloque de referencia para realizar una interpredicción de un bloque en una trama o imagen de vídeo subsiguiente.

[0192] De esta manera, el dispositivo de codificación 104 de la FIG. 7 representa un ejemplo de un codificador de vídeo configurado para obtener parámetros LIC, determinar adaptativamente tamaños de plantillas y/o seleccionar adaptativamente ponderaciones. El dispositivo de codificación 104 puede, por ejemplo, obtener parámetros LIC, determinar adaptativamente tamaños de plantillas y/o seleccionar adaptativamente conjuntos de ponderaciones como se describió anteriormente. Por ejemplo, el dispositivo de codificación 104 puede realizar cualquiera de las técnicas descritas en el presente documento, incluidos los procesos descritos anteriormente con respecto a la FIG. 5. En tales casos, el dispositivo de codificación 104 está configurado para generar archivos de formato de medios. En otros casos, algunas de las técnicas de esta divulgación pueden implementarse asimismo mediante el dispositivo de procesamiento posterior 57.

[0193] La FIG. 8 es un diagrama de bloques que ilustra un dispositivo de decodificación 112 de ejemplo. El dispositivo de decodificación 112 incluye una unidad de decodificación por entropía 80, una unidad de procesamiento de predicción 81, una unidad de cuantificación inversa 86, una unidad de procesamiento de transformación inversa 88, un sumador 90, una unidad de filtro 91 y una memoria de imágenes 92. La unidad de procesamiento de predicción 81 incluye una unidad de compensación de movimiento 82 y una unidad de procesamiento de intrapredicción 84. El dispositivo de decodificación 112 puede, en algunos ejemplos, realizar una pasada de decodificación que, en general, es recíproca a la pasada de codificación descrita con respecto al dispositivo de codificación 104 de la FIG. 7.

[0194] Durante el proceso de decodificación, el dispositivo de decodificación 112 recibe un flujo de bits de vídeo codificado, que representa los bloques de vídeo de un fragmento de vídeo codificado y elementos de sintaxis asociados, enviados por el dispositivo de codificación 104. En algunos modos de realización, el dispositivo de decodificación 112 puede recibir el flujo de bits de vídeo codificado desde el dispositivo de codificación 104. En algunos modos de realización, el dispositivo de decodificación 112 puede recibir el flujo de bits de vídeo codificado desde una entidad de red 79, tal como un servidor, un elemento de red sensible a los medios (MANE), un editor/empalmador de vídeo u otro dispositivo similar configurado para implementar una o más de las técnicas descritas anteriormente. La entidad de red 79 puede incluir o no el dispositivo de codificación 104. La entidad de red 79 puede implementar algunas de las técnicas descritas en esta divulgación, antes de que la entidad de red 79 transmita el flujo de bits de vídeo codificado al dispositivo de decodificación 112. En algunos sistemas de decodificación de vídeo, la entidad de red 79 y el dispositivo de decodificación 112 pueden formar parte de dispositivos independientes, mientras que, en otros casos, el mismo dispositivo que comprende el dispositivo de decodificación 112 puede realizar la funcionalidad descrita con respecto a la entidad de red 79.

[0195] La unidad de decodificación por entropía 80 del dispositivo de decodificación 112 decodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos de sintaxis. La unidad de decodificación por entropía 80 reenvía los vectores de movimiento y otros elementos de sintaxis a la unidad de procesamiento de predicción 81. El dispositivo de decodificación 112 puede recibir los elementos de sintaxis en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo. La unidad de

descodificación por entropía 80 puede procesar y analizar sintácticamente elementos de sintaxis de longitud fija y elementos de sintaxis de longitud variable en uno o más conjuntos de parámetros, tales como un VPS, SPS y PPS.

[0196] Cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de procesamiento de intrapredicción 84 de la unidad de procesamiento de predicción 81 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual basándose en un modo de intrapredicción señalizado y en datos de bloques previamente descodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento intercodificado (es decir, B, P o GPB), la unidad de compensación de movimiento 82 de la unidad de procesamiento de predicción 81 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y en otros elementos de sintaxis recibidos desde la unidad de descodificación por entropía 80. Los bloques predictivos se pueden producir a partir de una de las imágenes de referencia dentro de una lista de imágenes de referencia. El dispositivo de descodificación 112 puede construir las listas de tramas de referencia, la Lista 0 y la Lista 1, usando técnicas de construcción predeterminadas, basándose en las imágenes de referencia almacenadas en la memoria de imágenes 92.

[0197] La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando los vectores de movimiento y otros elementos de sintaxis, y usa la información de predicción para producir los bloques predictivos para el bloque de vídeo actual que se esté descodificando. Por ejemplo, la unidad de compensación de movimiento 82 puede usar uno o más elementos de sintaxis de un conjunto de parámetros para determinar un modo de predicción (por ejemplo, intrapredicción o interpredicción) usado para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de interpredicción (por ejemplo, un fragmento B, un fragmento P o un fragmento GPB), información de construcción para una o más listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, el estado de interpredicción para cada bloque de vídeo intercodificado del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

[0198] La unidad de compensación de movimiento 82 también puede realizar la interpolación en base a filtros de interpolación. La unidad de compensación de movimiento 82 puede usar filtros de interpolación como se usan mediante el dispositivo de codificación 104 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de los bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación usados por el dispositivo de codificación 104 a partir de los elementos de sintaxis recibidos y puede usar los filtros de interpolación para producir bloques predictivos.

[0199] La unidad de cuantificación inversa 86 cuantifica inversamente, o descuantifica, los coeficientes de transformación cuantificados proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el dispositivo de codificación 104 para cada bloque de vídeo del fragmento de vídeo para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que se debería aplicar. La unidad de procesamiento de transformación inversa 88 aplica una transformación inversa (por ejemplo una DCT inversa u otra transformación inversa adecuada), una transformación entera inversa o un proceso de transformación inversa conceptualmente similar, a los coeficientes de transformación para generar bloques residuales en el dominio de píxel.

[0200] Una vez que la unidad de compensación de movimiento 82 ha generado el bloque predictivo para el bloque de vídeo actual basándose en los vectores de movimiento y en otros elementos de sintaxis, el dispositivo de descodificación 112 forma un bloque de vídeo descodificado sumando los bloques residuales de la unidad de procesamiento de transformación inversa 88 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también pueden usarse filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad de vídeo. La unidad de filtro 91 pretende representar uno o más filtros de bucle tales como un filtro de desbloqueo, un filtro de bucle adaptativo (ALF) y un filtro de desviación adaptativa de muestras (SAO). Aunque la unidad de filtro 91 se muestra en la FIG. 8 siendo un filtro de bucle, en otras configuraciones la unidad de filtro 91 se puede implementar como un filtro de bucle posterior. Los bloques de vídeo descodificados en una trama o imagen dada se almacenan, a continuación, en la memoria de imágenes 92, que almacena imágenes de referencia usadas para la posterior compensación de movimiento. La memoria de imágenes 92 también almacena vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de destino de vídeo 122 mostrado en la FIG. 1.

[0201] De esta manera, el dispositivo de descodificación 112 de la FIG. 8 representa un ejemplo de un decodificador de vídeo configurado para obtener parámetros LIC, determinar adaptativamente tamaños de plantillas y/o seleccionar adaptativamente ponderaciones. El dispositivo de descodificación 112 puede, por ejemplo, obtener parámetros LIC, determinar adaptativamente tamaños de plantillas y/o seleccionar

adaptativamente conjuntos de ponderaciones como se describió anteriormente. Por ejemplo, el dispositivo de descodificación 112 puede realizar cualquiera de las técnicas descritas en el presente documento, incluidos los procesos descritos anteriormente con respecto a la FIG. 6. En tales casos, el dispositivo de descodificación 112 está configurado para analizar archivos de formato de medios.

[0202] En la descripción anterior, se describen aspectos de la aplicación con referencia a modos de realización específicos de la misma, pero los expertos en la técnica reconocerán que la invención no está limitada a los mismos. Por lo tanto, si bien los modos de realización ilustrativos de la solicitud se han descrito en detalle en el presente documento, debe entenderse que los conceptos de la invención pueden incorporarse y emplearse de otro modo, y que las reivindicaciones adjuntas están destinadas a interpretarse para incluir tales variaciones, excepto de la forma en la que están limitadas por el estado de la técnica anterior. Diversas características y aspectos de la invención descrita anteriormente se pueden usar individual o conjuntamente. Se pueden utilizar otros modos de realización en cualquier número de entornos y aplicaciones más allá de los descritos en el presente documento sin apartarse del alcance de la invención tal como se define en las reivindicaciones adjuntas. La memoria descriptiva y los dibujos, en consecuencia, deben considerarse en un sentido ilustrativo, en lugar de en un sentido restrictivo. Para fines de ilustración, los procedimientos se han descrito en un orden particular. Debe apreciarse que en modos de realización alternativos, los procedimientos se pueden realizar en un orden diferente al descrito.

[0203] Cuando los componentes se describen como "configurados para" realizar determinadas operaciones, dicha configuración se puede lograr, por ejemplo, mediante el diseño de circuitos electrónicos u otro hardware para realizar la operación, mediante la programación de circuitos electrónicos programables (por ejemplo, microprocesadores u otros circuitos electrónicos adecuados) para realizar la operación, o cualquier combinación de lo anterior.

[0204] Los diversos bloques lógicos, módulos, circuitos y etapas de algoritmo ilustrativos descritos en relación con los modos de realización divulgados en el presente documento pueden implementarse como hardware electrónico, software informático, firmware o combinaciones de los mismos. Para ilustrar claramente esta intercambiabilidad de hardware y software, anteriormente se han descrito en general diversos componentes, bloques, módulos, circuitos y etapas ilustrativos en lo que respecta a su funcionalidad. Que dicha funcionalidad se implemente como hardware o software depende de la aplicación particular y de las restricciones de diseño impuestas al sistema global. Los expertos en la técnica pueden implementar la funcionalidad descrita de formas distintas para cada aplicación particular, pero no debería interpretarse que dichas decisiones de implementación suponen apartarse del alcance de la presente invención.

[0205] Las técnicas descritas en el presente documento también se pueden implementar en hardware electrónico, software informático, firmware o en cualquier combinación de los mismos. Dichas técnicas se pueden implementar en cualquiera entre una variedad de dispositivos tales como ordenadores de propósito general, equipos manuales de dispositivos de comunicación inalámbrica o dispositivos de circuitos integrados que tienen múltiples usos, incluyendo su aplicación en equipos manuales de dispositivos de comunicación inalámbrica y otros dispositivos. Todos los rasgos característicos descritos como módulos o componentes se pueden implementar juntos en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. Si se implementan en software, las técnicas se pueden realizar, al menos en parte, por un medio de almacenamiento de datos legible por ordenador que comprenda código de programa que incluya instrucciones que, cuando se ejecuten, realicen uno o más de los procedimientos descritos anteriormente. El medio de almacenamiento de datos legible por ordenador puede formar parte de un producto de programa informático, que puede incluir materiales de embalaje. El medio legible por ordenador puede comprender memoria o medios de almacenamiento de datos, tales como memoria de acceso aleatorio (RAM), tal como memoria de acceso aleatorio dinámica síncrona (SDRAM), memoria de solo lectura (ROM), memoria de acceso aleatorio no volátil (NVRAM), memoria de solo lectura programable y borrrable eléctricamente (EEPROM), memoria FLASH, medios de almacenamiento de datos magnéticos u ópticos, y similares. Las técnicas se pueden realizar, adicionalmente o de forma alternativa, al menos en parte, por un medio de comunicación legible por ordenador que transporta o comunica código de programa en forma de instrucciones o estructuras de datos y al que se puede acceder, leer y/o ejecutar por medio de un ordenador, tales como señales u ondas propagadas.

[0206] El código de programa se puede ejecutar por un procesador, que puede incluir uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables en el terreno (FPGA) u otros circuitos lógicos equivalentes, integrados o discretos. Un procesador de este tipo puede estar configurado para realizar cualquiera de las técnicas descritas en la presente divulgación. Un procesador de propósito general puede ser un microprocesador, pero, de forma alternativa, el procesador puede ser cualquier procesador, controlador, microcontrolador o máquina de estados convencional. Un procesador también se puede implementar como una combinación de dispositivos informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo. Por consiguiente, el término "procesador", como se usa en el presente documento, se puede referir a cualquier estructura anterior, cualquier combinación de la estructura

anterior, o cualquier otra estructura o aparato adecuados para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de software o módulos de hardware dedicados configurados para la codificación y la decodificación, o incorporados en un codificador-descodificador de vídeo combinado (CODEC).

REIVINDICACIONES

1. Un procedimiento de procesamiento de datos de vídeo, comprendiendo el procedimiento:

5 obtener los datos de vídeo; y

generar un archivo de formato de medios, formateado de acuerdo con el formato de archivo de medios de base ISO usando un esquema restringido limitado para permitir que se incluyan múltiples cajas de tipo de esquema en una caja de información de esquema restringido, incluyendo el archivo de formato de medios los datos de vídeo;

generar, para una entrada de muestras del archivo de formato de medios, una caja de información de esquema restringido, estando asociada la entrada de muestras con una o más muestras de los datos de vídeo,

generar, para la caja de información de esquema restringido, una pluralidad de cajas de tipo de esquema que identifican una pluralidad de tipos de esquema restringido requeridos para procesar una o más muestras de los datos de vídeo asociados con la entrada de muestras.

2. Un procedimiento de procesamiento de uno o más archivos de formato de medios, que comprende:

obtener un archivo de formato de medios, formateado de acuerdo con el formato de archivo de medios de base ISO usando un esquema restringido limitado para permitir que se incluyan múltiples cajas de tipo de esquema en una caja de información de esquema restringido, incluyendo el archivo de formato de medios datos de vídeo y una caja de información de esquema restringido para un entrada de muestras del archivo de formato de medios, estando asociada la entrada de muestras con una o más muestras de los datos de vídeo, incluyendo la caja de información de esquema restringido una pluralidad de cajas de tipo de esquema, en el que la pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido requeridos para procesar los datos de vídeo;

analizar el archivo de formato de medios; y

procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquema restringido identificados por la pluralidad de cajas de tipo de esquema.

3. El procedimiento según la reivindicación 1 o 2, en el que cada caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica un tipo de esquema restringido diferente requerido para procesar los datos de vídeo.

4. El procedimiento según la reivindicación 1 o 2, en el que una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema identifica una versión de un tipo de esquema restringido identificado por la caja de tipo de esquema.

5. El procedimiento según la reivindicación 1 o 2, en el que una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema incluye un identificador que identifica una fuente de datos necesarios para implementar un tipo de esquema restringido identificado por la caja de tipo de esquema.

6. El procedimiento según la reivindicación 1 o 2, en el que la caja de información de esquema restringido incluye, además, una pluralidad de cajas de información de esquema, y en el que una caja de información de esquema de la pluralidad de cajas de información de esquema incluye datos necesarios para un tipo de esquema restringido identificado por una caja de tipo de esquema de la pluralidad de cajas de tipo de esquema.

7. El procedimiento según la reivindicación 1, que comprende, además, transmitir el archivo de formato de medios.

8. El procedimiento según la reivindicación 1, que comprende, además, almacenar el archivo de formato de medios.

9. El procedimiento según la reivindicación 2, que comprende, además, decodificar los datos de vídeo.

10. El procedimiento según la reivindicación 2, que comprende, además, mostrar los datos de vídeo decodificados.

11. Un medio legible por ordenador no transitorio que almacena instrucciones que, cuando son ejecutadas por uno o más procesadores, hacen que el uno o más procesadores realicen el procedimiento de cualquiera de las reivindicaciones 1 a 10.

12. Un aparato para procesar datos de vídeo, que comprende:

una memoria configurada para almacenar los datos de vídeo; y

un procesador configurado para:

obtener los datos de vídeo; y

generar un archivo de formato de medios, formateado de acuerdo con el formato de archivo de medios de base ISO usando un esquema restringido limitado para permitir que se incluyan múltiples cajas de tipo de esquema en una caja de información de esquema restringido, incluyendo el archivo de formato de medios los datos de vídeo,

generar, para una entrada de muestras del archivo de formato de medios, una caja de información de esquema restringido, estando asociada la entrada de muestras con una o más muestras de los datos de vídeo,

generar, para la caja de información de esquema restringido, una pluralidad de cajas de tipo de esquema que identifican una pluralidad de tipos de esquema restringido requeridos para procesar una o más muestras de los datos de vídeo asociados con la entrada de muestras.

13. El aparato según la reivindicación 12, en el que el aparato comprende un dispositivo móvil con una cámara para capturar imágenes.

14. Un aparato para procesar uno o más archivos de formato de medios, que comprende:

una memoria configurada para almacenar datos de vídeo y uno o más archivos de formato de medios; y

un procesador configurado para:

obtener un archivo de formato de medios, formateado de acuerdo con el formato de archivo de medios de base ISO usando un esquema restringido limitado para permitir que se incluyan múltiples cajas de tipo de esquema en una caja de información de esquema restringido, incluyendo el archivo de formato de medios los datos de vídeo y una caja de información de esquema restringido para una entrada de muestras del archivo de formato de medios, estando asociada la entrada de muestras con una o más muestras de los datos de vídeo, incluyendo la caja de información de esquema restringido una pluralidad de cajas de tipo de esquema, en el que la pluralidad de cajas de tipo de esquema identifica una pluralidad de tipos de esquema restringido requeridos para procesar los datos de vídeo;

analizar el archivo de formato de medios; y

procesar los datos de vídeo de acuerdo con la pluralidad de tipos de esquemas restringidos identificados por la pluralidad de cajas de tipo de esquema.

15. El aparato según la reivindicación 14, que comprende, además:

un decodificador para decodificar los datos de vídeo.

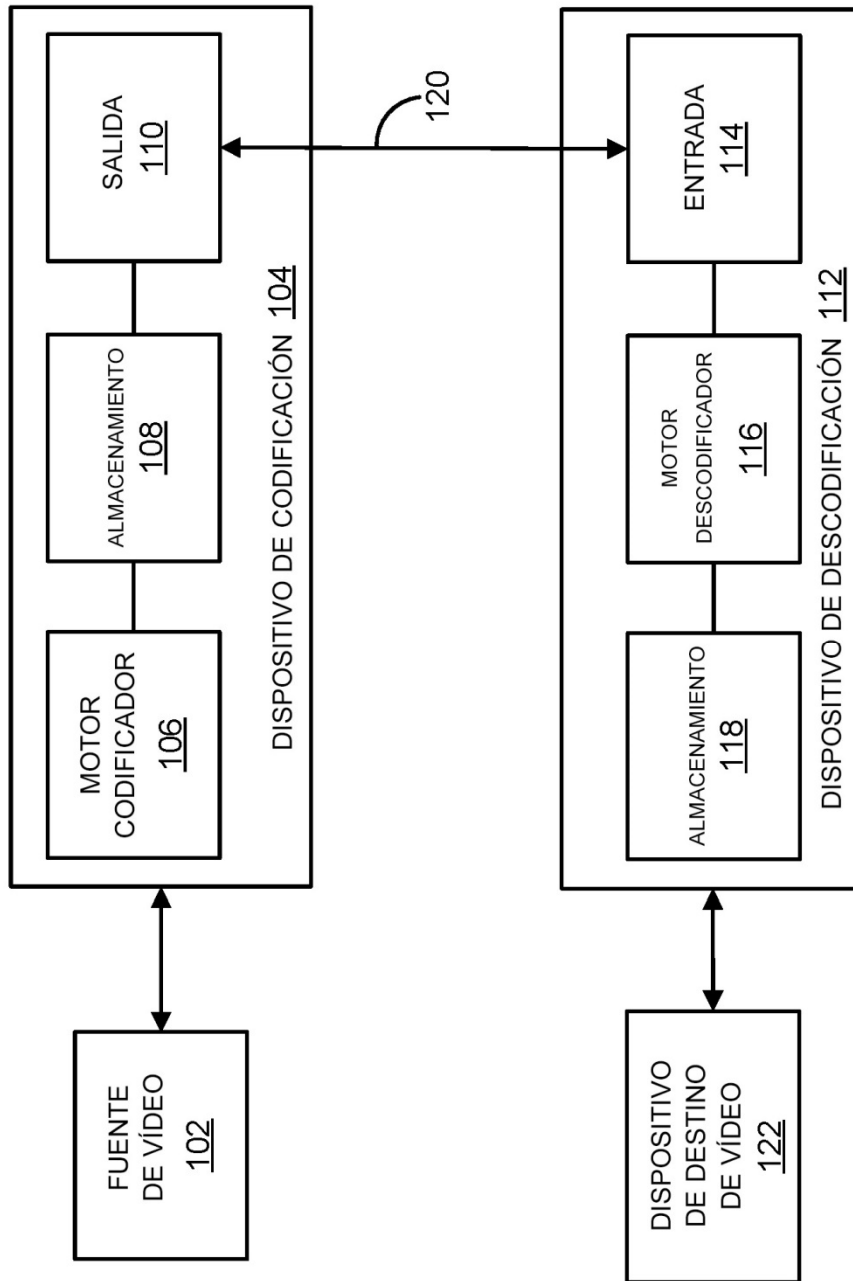


FIG. 1

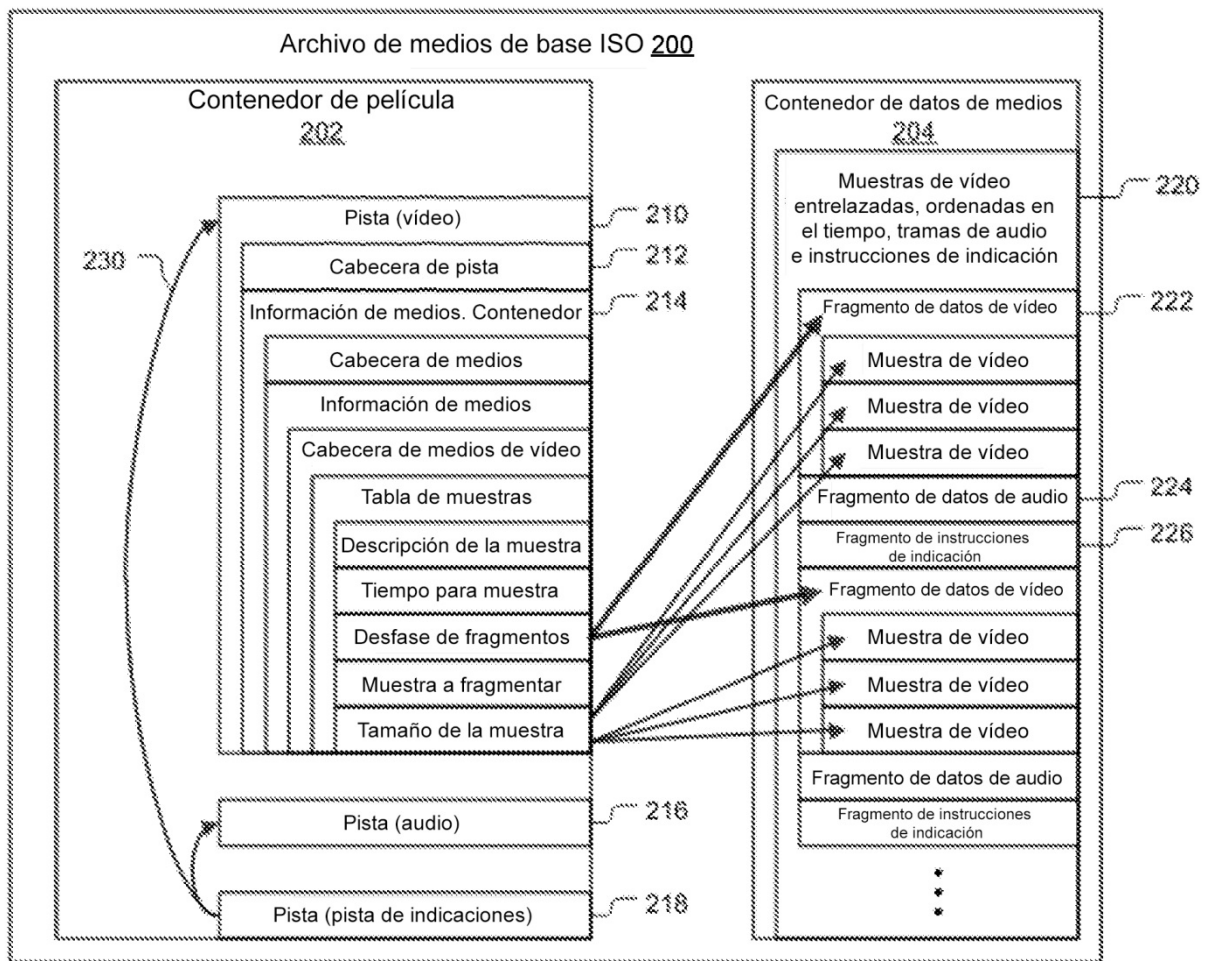


FIG. 2

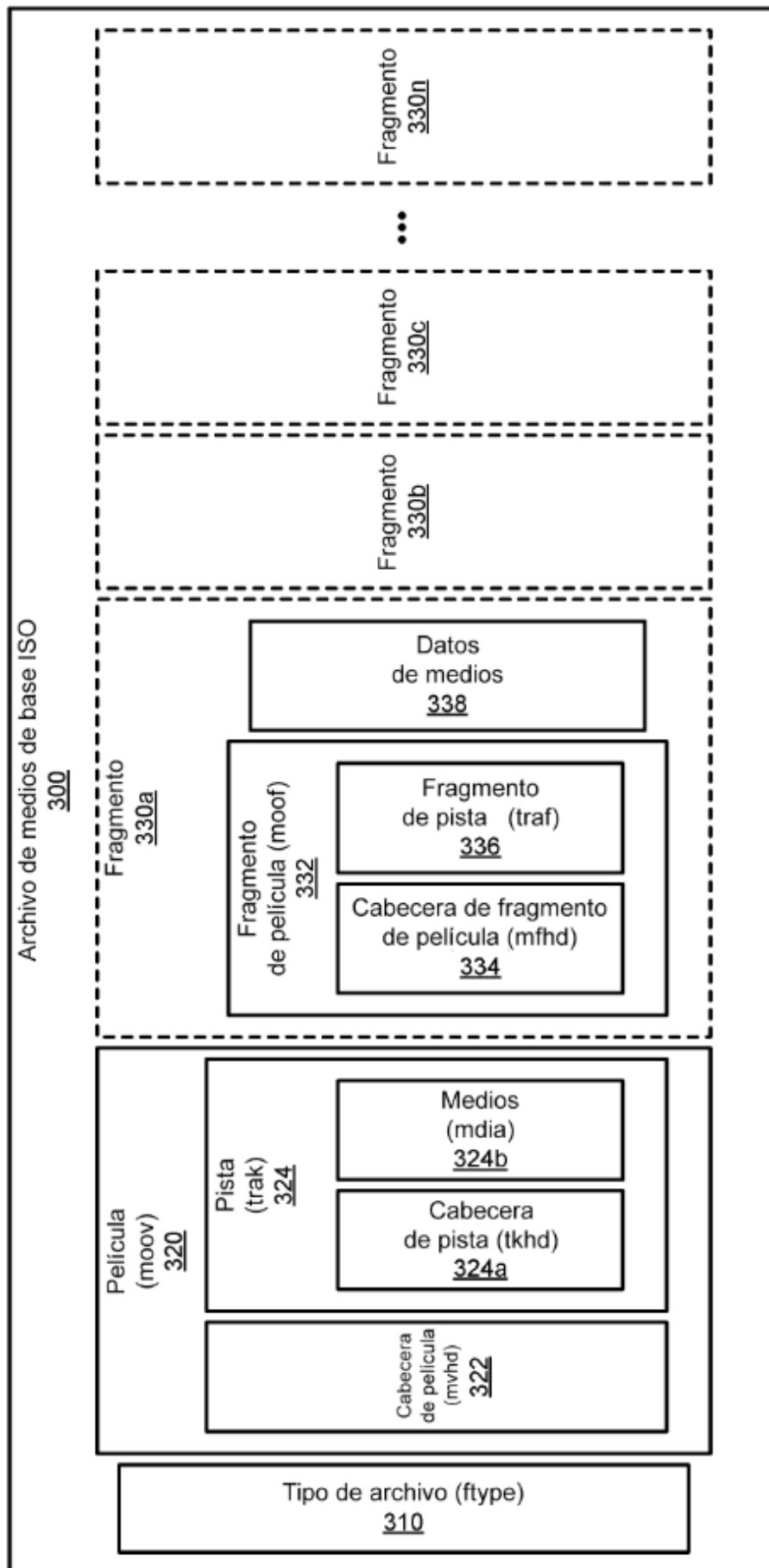


FIG. 3

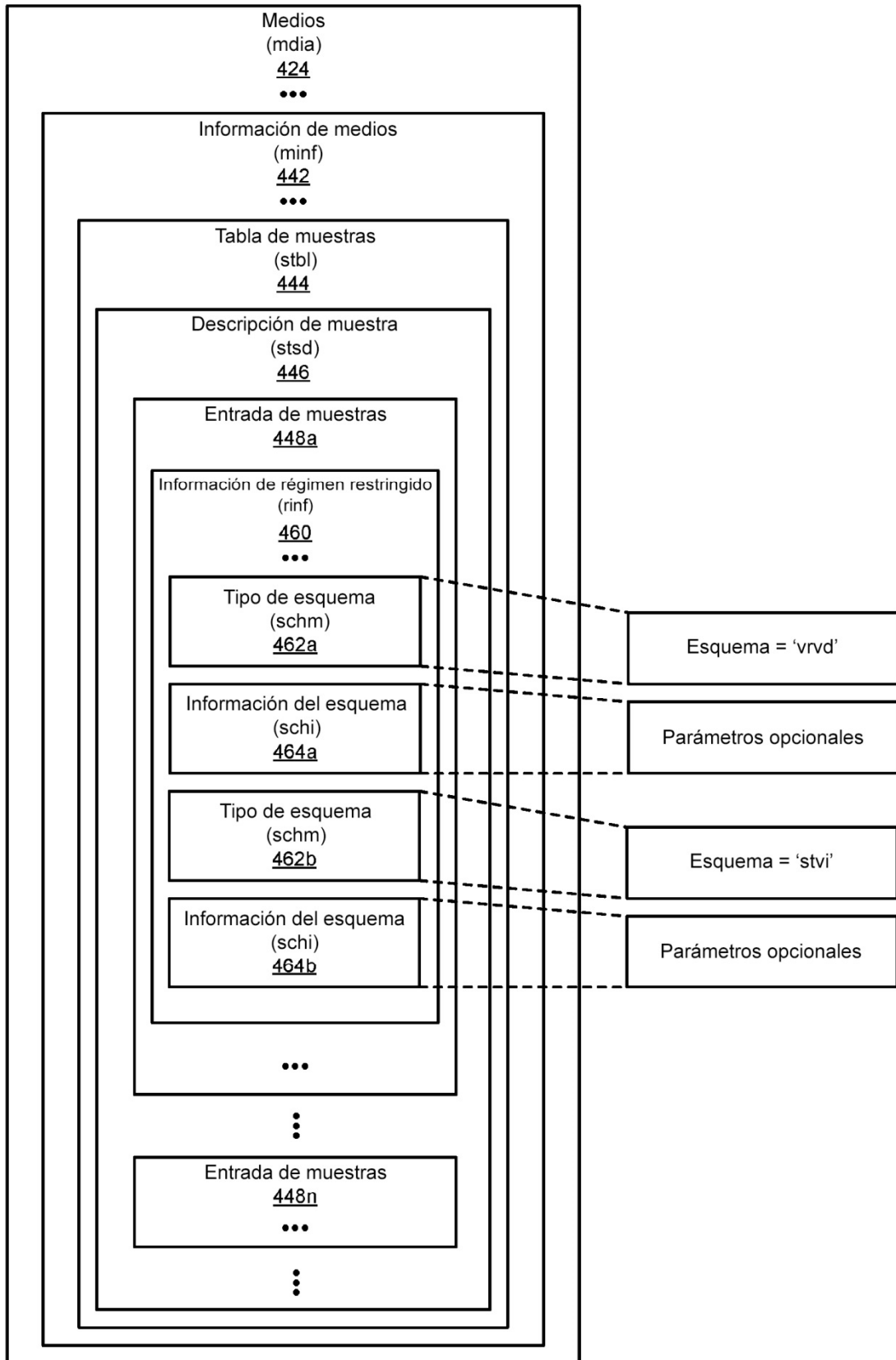


FIG. 4

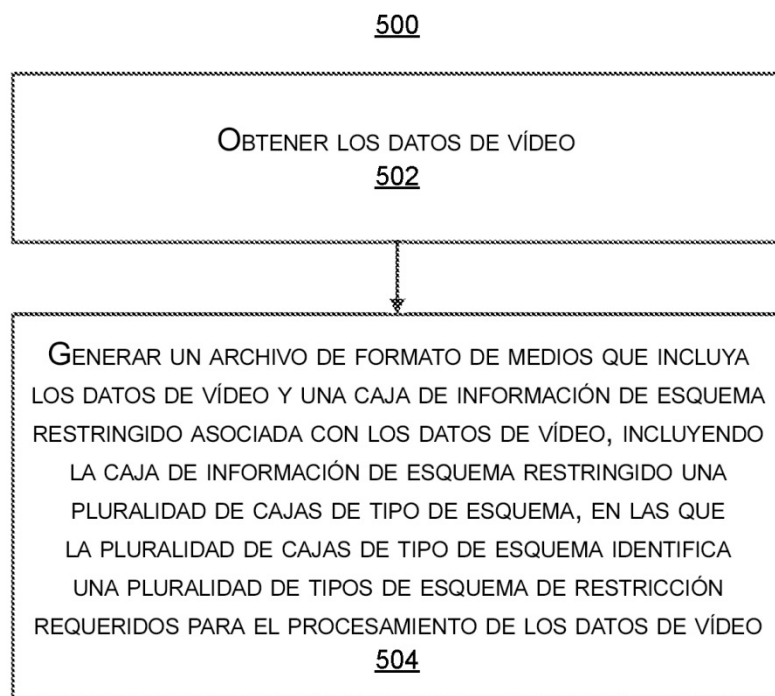


FIG. 5

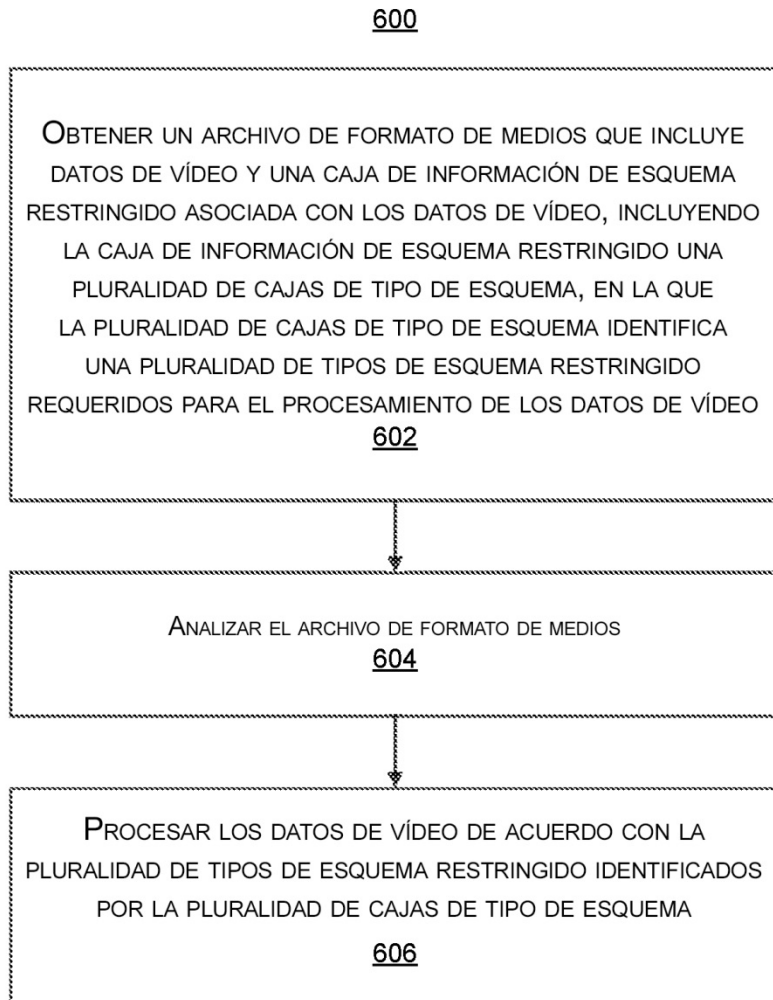


FIG. 6

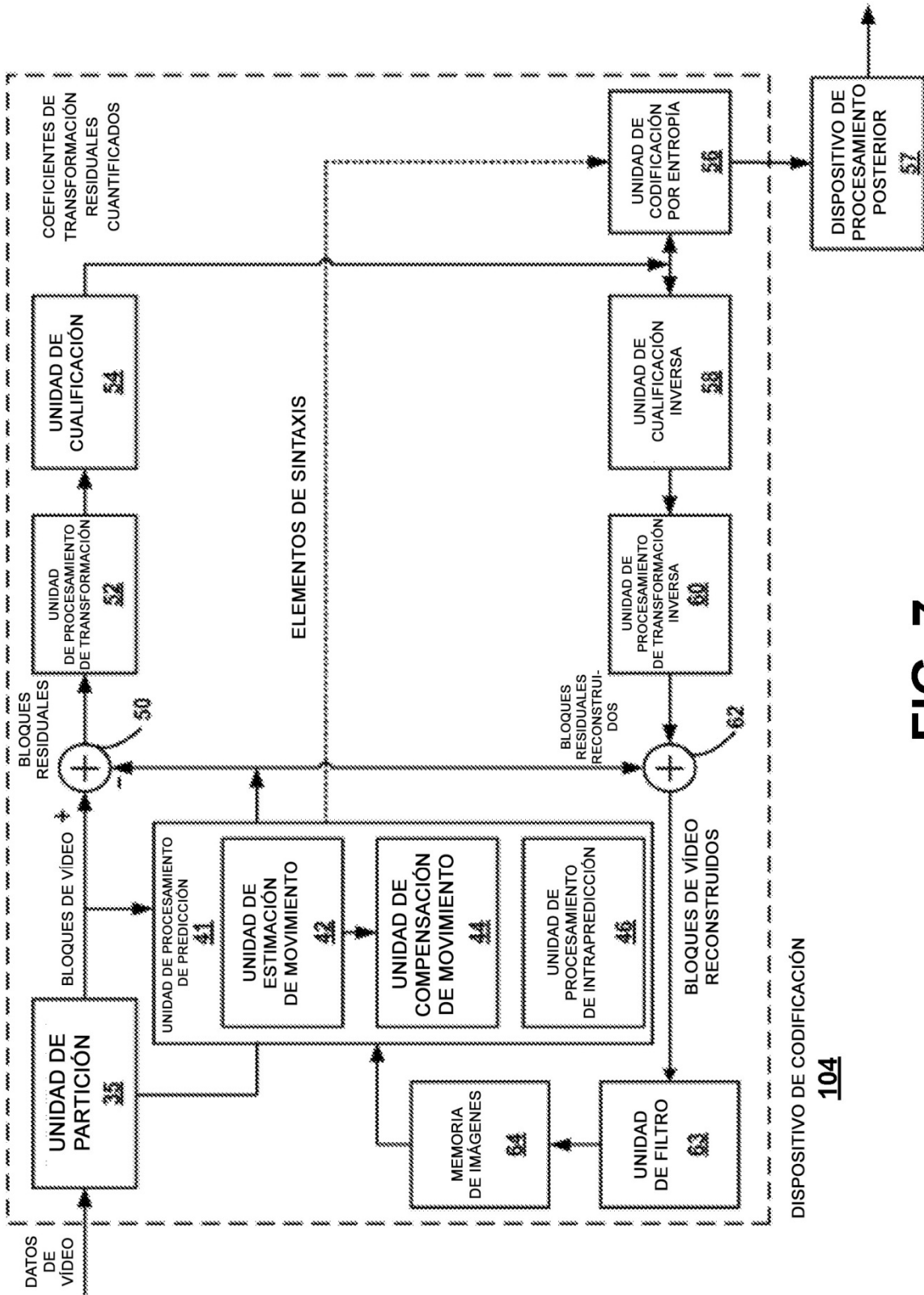


FIG. 7

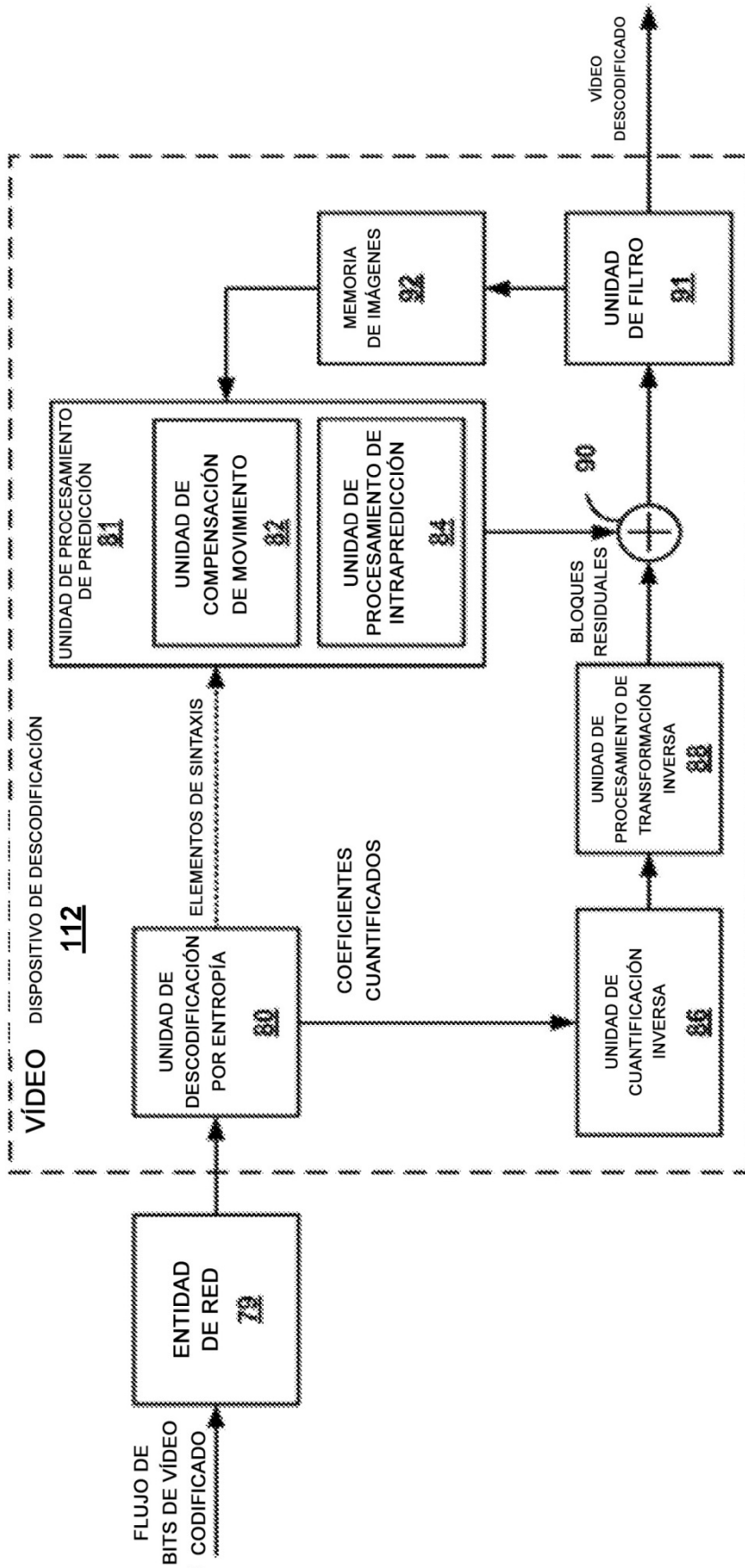


FIG. 8